

(19)



(11)

**EP 4 041 048 B1**

(12)

**EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:

**01.01.2025 Bulletin 2025/01**

(21) Application number: **20874325.2**

(22) Date of filing: **07.10.2020**

(51) International Patent Classification (IPC):

**A61B 1/00** (2006.01)      **A61B 1/04** (2006.01)  
**A61B 1/05** (2006.01)      **G06T 7/70** (2017.01)  
**G06T 7/73** (2017.01)      **G06T 7/80** (2017.01)

(52) Cooperative Patent Classification (CPC):

**A61B 1/00179; A61B 1/000094; A61B 1/000096;**  
**A61B 1/00005; A61B 1/00183; A61B 1/00188;**  
**G02B 23/2484; G06T 7/80; G06T 15/205;**  
**G06T 2210/41**

(86) International application number:

**PCT/US2020/054640**

(87) International publication number:

**WO 2021/071991 (15.04.2021 Gazette 2021/15)**

**(54) SYSTEMS AND METHODS FOR CHANGING THE DIRECTION OF VIEW DURING VIDEO GUIDED CLINICAL PROCEDURES USING REAL-TIME IMAGE PROCESSING**

SYSTEME UND VERFAHREN ZUR ÄNDERUNG DER BLICKRICHTUNG BEI VIDEOGESTEUERTEN KLINISCHEN VERFAHREN MITTELS ECHTZEIT-BILDVERARBEITUNG

SYSTÈMES ET PROCÉDÉS POUR MODIFIER LA DIRECTION DE VISION PENDANT DES INTERVENTIONS CLINIQUES GUIDÉES PAR VIDÉO À L'AIDE D'UN TRAITEMENT D'IMAGE EN TEMPS RÉEL

(84) Designated Contracting States:

**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**

(30) Priority: **07.10.2019 US 201962911950 P**

**07.10.2019 US 201962911986 P**

(43) Date of publication of application:

**17.08.2022 Bulletin 2022/33**

(73) Proprietor: **S&N Orion Prime, S.A.**

**3030-199 Coimbra (PT)**

(72) Inventors:

- **DE ALMEIDA BARRETO, João Pedro**  
**3000-031 Coimbra (PT)**
- **DOS SANTOS RAPOSO, Carolina**  
**3030-502 Coimbra (PT)**
- **ALMEIDA ANTUNES, Michel Goncalves**  
**3020-171 Coimbra (PT)**
- **TEIXEIRA, Rui Jorge Melo**  
**3460-613 Tondela (PT)**

(74) Representative: **White, Andrew John et al**

**Mathys & Squire**  
**The Shard**  
**32 London Bridge Street**  
**London SE1 9SG (GB)**

(56) References cited:

**US-A1- 2007 236 514      US-A1- 2014 285 676**  
**US-A1- 2015 065 793      US-A1- 2016 048 953**

- **"Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003", vol. 2879, 15 November 2003, SPRINGER BERLIN HEIDELBERG, Berlin, Heidelberg, ISBN: 978-3-540-39903-2, ISSN: 0302-9743, article YAMAGUCHI TETSUZO ET AL: "Camera Model and Calibration Procedure for Oblique-Viewing Endoscope", pages: 373 - 381, XP093086524, DOI: 10.1007/978-3-540-39903-2\_46**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

**EP 4 041 048 B1**

- GIRARD FANTIN ET AL: "Active self-calibration of thoracoscopic images for assisted minimally invasive spinal surgery", LASER-BASED MICRO-AND NANOPACKAGING AND ASSEMBLY II, vol. 7625, 4 March 2010 (2010-03-04), pages 76250Q - 1, XP093086547, ISSN: 0277-786X, Retrieved from the Internet <URL:<http://dx.doi.org/10.1117/12.843985>> DOI: 10.1117/12.843985

**Description****CROSS-REFERENCE TO RELATED APPLICATIONS**

5 **[0001]** This application claims priority to and the benefit of U.S. Provisional Patent Application No. 62/911,986, filed October 7, 2019 ("986 application") and U.S. Application No. 62/911,950, filed on October 7, 2019 ("950 application").

**FIELD**

10 **[0002]** The disclosure generally relates to the fields of computer vision and image processing, and in particular, but not by way of limitation, the presented disclosed embodiments can be used for enhanced visualization in video guided minimally invasive clinical procedures of surgery and diagnosis, such as arthroscopy, laparoscopy or endoscopy, for the purposes of changing the direction of view of the surgical camera with an endoscopic lens, in which case the system renders the view that would be acquired by a physical scope with a different lens cut from the one that is being effectively  
15 used, or performing zoom along an arbitrary viewing direction without decreasing the image field-of-view or losing image contents.

**BACKGROUND**

20 **[0003]** In video-guided procedures, such as arthroscopy or laparoscopy, the anatomical cavity of interest is accessed through small incisions designated as surgical ports. One of these ports gives access to a video camera equipped with a rigid endoscope that enables the surgeon to visualize the interior of the cavity for the purpose of guidance during procedures of surgery or diagnosis. The rigid endoscope is an elongated tubular structure that is inserted into the body cavity. The endoscope typically comprises an objective lens at the distal end, an image-forwarding system, such as a  
25 series of spaced-apart lenses, and an ocular lens at the proximal end where a camera means, such as a charge coupled device (CCD) chip, is usually mounted. The image-forwarding system serves to forward the image from distal to proximal end. The camera receives this image and produces a signal for a video display to be observed by the surgeon.

**[0004]** The rigid endoscope, also referred to in here as lens scope or endoscopic lens, is used in different specialties and, depending on its features and application field, it can be alternatively named as arthroscope (orthopedics), laparoscope (abdominal surgery), cystoscope (bladder), sinuscope (ENT), neuroscope, etc. The endoscopic camera, that results from  
30 combining a rigid endoscope with a camera means, has specific features that are uncommon in conventional cameras. The optics are usually exchangeable for the purpose of easy sterilization, with the endoscope being assembled in the camera head by the surgeon before starting the procedure. The ocular lens (or eye-piece) in the proximal end of the endoscope is assembled to the camera using a connector that typically allows the user to rotate the scope with respect to  
35 the camera head by an angle  $\alpha$ . This is a rotation in azimuth around a longitudinal axis of the endoscope, henceforth referred to as mechanical axis, that is close, but not necessarily coincident, with the symmetry axis of the tubular structure.

**[0005]** The rigid endoscope typically has a Field Stop Mask (FSM) somewhere along the image forwarding system that causes the acquired image to have meaningful content in a circular region which is surrounded by a black frame (which black frame is illustrated in diagonal lines in the figures of this application). The design of the FSM is such that there is  
40 usually a mark in the circular boundary defined by the image-frame transition that enables the surgeon to infer the angle in azimuth  $\alpha$ . This mark in the periphery of the circular image will be henceforth referred to as the notch.

**[0006]** In addition, the lens in the distal end of the endoscope is often mounted in such a way that the optical axis of the entire lens system has an angular offset from the longitudinal or mechanical axis of the scope. This angle  $\beta$  between optical and mechanical axes is usually referred to as the lens cut of the endoscope. If the lens cut is different from zero, then the  
45 optical and mechanical axes are not aligned and the surgeon can vary the direction of view in azimuth by rotating the scope around the mechanical axis without having to move the camera head.

**[0007]** A variable Direction of View (DoV) enables the surgeon to change the endoscopic viewing direction without having to change the position of the endoscope itself. This is useful to see structures which are beside or behind the tip of the endoscope whenever the endoscope shaft cannot be easily moved because of anatomical constraints or constraints  
50 imposed by other surgical instruments in the operative field. Thus, variable DoV is definitely a desirable feature that affords the surgeon greater flexibility in his/her procedural approach. As discussed, the rigid endoscopes currently used in clinic empower the surgeon with the possibility of varying the DoV by performing a rotation in azimuth of the lens scope, which causes the optical axis to describe a cone in space with apex close to the projection center O (the cone of DoV). The half-angle of this cone is defined by the lens cut  $\beta$  that is fixed and known *a priori* by the surgeon. Surgeons rely heavily on the  
55 prior knowledge of the lens cut of a particular endoscope to dependably know what the anatomy should look like.

**[0008]** There are endoscopes with different angular offsets from the longitudinal or mechanical axis, with the most commonly used lens cuts being  $0^\circ$ ,  $30^\circ$ ,  $45^\circ$  and  $70^\circ$ . Surgical procedures typically require endoscopes of most of these cut angles with specific emphasis on one of them. For example, in knee arthroscopy the preference goes for a lens cut of  $30^\circ$

that provides both a good forward view and a certain amount of lateral viewing, while in hip arthroscopy the lens cut of choice is usually 70° because the anatomic cavity is much narrower. In a similar manner, the preference in laparoscopy can vary between a lens cut of 0° (forward looking scope) or 30°. However, and despite the emphasis in a particular lens cut angle, most procedures can benefit from different lateral and partial retro-viewing angles at different moments or steps of the surgery. Unfortunately, and since a lateral change in DoV requires to physically exchange the endoscope to one with a different lens cut, the surgeons rarely do it in practice even when they realize that such change can improve visualization to accomplish the task at hands. Changing the endoscope in the middle of the procedure is cumbersome (both light and camera cables have to be disconnected and reconnected), time consuming, and sometimes dangerous. For example, the insertion of off-angle endoscopes can be dangerous because they are not looking in the direction they are being inserted, and neurosurgeons often refrain from using endoscopes with 45° or 70° lens cuts because they are afraid of blindly thrusting the endoscope into delicate tissue.

**[0009]** In summary, the rigid endoscopes commonly used in medicine enable the surgeon to vary the DoV in azimuth (angle  $\alpha$ ) but not in inclination or elevation (angle  $\beta$ ), in which case the optical axis is constrained to move in a cone with symmetry axis aligned with the mechanical axis and angle defined by the lens cut  $\beta$  (the cone of DoV). It follows from the discussion that unconstrained variation of DoV would be highly beneficial namely by allowing the surgeon to change the lateral viewing angle (the lens cut  $\beta$ ) without having to physically change the endoscope in the mid procedure. Several specially designed endoscopes, that allow angular shift of the direction of view both in azimuth and in inclination, have been disclosed (US2016/02.09A1, US20050177026A1, US2012/0078049, US9,307,892). Unfortunately, these endoscopes use a movable image sensor or optical element at the distal end to change the lens cut  $\beta$ . Because of these moving parts, fabricating these scopes is complicated and costly, and such scopes are less robust than traditional fixed lens cut scopes. Also, they often deliver inferior illumination and image quality.

**[0010]** An alternative manner for accomplishing unconstrained variation of DoV is to employ computational means to process the endoscopic image or video. It is well known in the field of computer vision that a change in the DoV corresponds to a rotation of the camera around an axis going through its projection center, and that the virtual image that would be acquired by such rotated camera can be rendered by warping the real source image by a suitable homographic map often referred to as a collineation of the plane at infinity. The method is used in Patent Applications US005313306A, US 20154/0065799A1, and EP3130276A1 that disclosed rigid endoscopes with wide angle lenses for hemispherical Field-of-View (FoV), that are combined with computer means to render images with a variable DoV that cover selected parts of the FoV. The result are camera systems with controllable pan-and-tilt orientation where the user commands the viewing direction both in X (azimuth) and Y (elevation) using electronic means.

**[0011]** Document US2014285676A1 (BARRETO et Al) 25.09.2014, discloses a calibration method where the intrinsic parameters are updated at each rotation by using the principal point, the boundary contour and the triangular mark (i.e. notch).

**[0012]** A problem of these systems is that in the context of endoscopy and video-guided clinical procedures the surgeon is used to changing the viewing direction in azimuth by simply rotating the lens scope with respect to the camera-head. Thus, moving the control from mechanical to electronic can be challenging, require surgical re-training, and become a major obstacle for broad adoption.

**[0013]** What is desired, therefore, is an endoscopy system with a DoV that can vary both in azimuth  $\alpha$  and inclination  $\beta$ , where the variation in azimuth is accomplished by the standard mechanical means, i.e. by rotating the rigid endoscope with respect to the camera head, and the variation in inclination is accomplished by electronic or computer means, i.e. by using real-time processing to render images with the targeted viewing direction.

**[0014]** It is still further desired to provide an endoscopy system that operates in the standard manner but where the user has the possibility to electronically switch between rigid endoscopes with different lens cuts for optimum visualization conditions at all times.

**[0015]** Additionally, it is desired to provide an endoscopy system where the user can control the amount of image radial distortion to either work with distortion free images for optimal perception of scene depth, or to select a Region of Interest by varying the DoV and magnify it while preserving the FoV and maintaining all visible contents in image.

## SUMMARY

**[0016]** The invention is set out in the appended set of claims. Systems and methods that use real-time image processing to change the direction of view during video-guided clinical procedures of surgery and diagnosis, while allowing the surgeon to physically rotate the lens scope with respect to the camera-head as is current common practice in different surgical procedures such as arthroscopy, laparoscopy and endoscopy.

**[0017]** The presently disclosed embodiments disclose a software-based system that enables to change the viewing direction of a particular scope by an arbitrary amount  $\gamma$  such that the cutting angle of the lens (or lens cut) virtually switches from  $\beta$  to  $\beta + \gamma$ . This is accomplished in a seamless, fully automatic manner that is realistic for the user that can physically rotate the endoscope in azimuth around its longitudinal axis during operation.

[0018] A software-based system processes the images and video acquired by an endoscopy camera equipped with a lens with a wide Field of View (FoV) to empower the surgeon with electronic switch between two or more virtual endoscopes with different lens cuts.

[0019] In addition, the presently disclosed embodiments disclose a method that, given the desired lens cuts  $\hat{\beta}_1$  and  $\hat{\beta}_2$  and FoVs  $\hat{\Theta}_1$  and  $\hat{\Theta}_2$  of the two targeted virtual endoscopes, determines the lens cut  $\beta$  and FoV  $\Theta$  of the source camera such that the acquired images and video are well suited to be used as input in the disclosed image processing methods for rendering realistic virtual images and video.

[0020] The presently disclosed embodiments disclose a method that, given arbitrary pre-sets for the virtual camera, not only in terms of desired shift in lens cut  $\gamma$ , but also in terms of image resolution  $m \times n$ , field-of-view  $\hat{\Theta}$ , and amount of image distortion  $\xi$ , assures that the corresponding virtual view is correctly rendered without top or bottom regions without content (empty regions) by automatically adjusting the focal length  $\hat{f}$  and position of principal point  $\hat{\mathbf{O}}$  at each frame time instant.

[0021] The presently disclosed embodiments disclose a method that enables to control the amount of image radial distortion to either work with distortion free images for optimal perception of scene depth, or to select a Region of Interest by performing image zoom along arbitrary viewing directions by either changing the amount of radial distortion  $\hat{\xi}$ , in which case the zoom is accomplished without loss in the field of view (FoV), or by changing the focal length  $\hat{f}$ .

## BRIEF DESCRIPTION OF THE DRAWING

[0022] For a more complete understanding of the present disclosure, reference is made to the following detailed description of exemplary embodiments considered in conjunction with the accompanying drawings.

FIG. 1 is an embodiment of an endoscopic camera comprising an endoscope and a camera head, and showing all its constituent parts. The canonical image that is formed by the endoscope that contains a Field Stop Mask (FSM) in its image forwarding system is depicted, as well as the image in pixels obtained after the application of the intrinsic parameters by the sensor in the camera head.

FIG. 2 is an embodiment of an endoscope, showing two different directions of viewing (DoV) that give rise to two distinct cones of DoV. Each cone corresponds to the trajectory described by the optical axis while the endoscope rotates around the mechanical axis.

FIG. 3 illustrates the image warping process between a source and a target image, with the color of the pixels in the target image being obtained using an interpolation scheme. The warping function is the composition of the source and target camera models  $c_s$  and  $c_t$  with a motion model  $m$ .

FIG. 4 schematizes the process of obtaining camera model  $c$ . A 3D point  $\mathbf{X}$  is projected onto the image plane according to the pinhole model, yielding a 2D point  $\mathbf{x}$  in the canonical image. Afterwards, distortion  $d$  is an operation performed in the canonical image that applies distortion to point  $\mathbf{x}$ , yielding point  $\mathbf{x}'$ . As a final step, the application of the intrinsic parameters by function  $k$  converts points in millimeters  $\mathbf{x}'$  into points  $\mathbf{u}$  in the image in pixels. The camera model  $c$  is a function applied in 2D space that transforms points  $\mathbf{x}$  in the canonical image into points  $\mathbf{u}$  in the image in pixels,  $\mathbf{u} = c(\mathbf{x}; f, \mathbf{O}, \xi)$ , and is obtained by the composition of the distortion  $d$  and the intrinsics  $k$ :  $c = k \circ d$ .

FIG. 5 illustrates the effect of the scope rotation in the location of the boundary and the principal point  $\mathbf{O}$ . The scope rotates around the mechanical axis that intersects the image plane in the rotation center  $\mathbf{Q}$ . The FSM is projected onto the image plane as a circle with center  $\mathbf{C}$ , and the mark in the FSM is projected onto a point  $\mathbf{P}$ , referred to as notch. When the scope rotates by an angle  $\delta_i$ , the circle center  $\mathbf{C}$ , notch  $\mathbf{P}$  and principal point  $\mathbf{O}$  also rotate by that same amount  $\delta_i$ , around  $\mathbf{Q}$ .

FIG. 6 illustrates the relevant entities in the estimation of the motion model  $m$ , which is a 3D rotation by an angle  $\gamma = \hat{\beta} - \beta$  around axis  $\vec{\mathbf{n}}_i$ . This rotation is the virtual motion the real camera with lens cut  $\beta$  must perform in order to virtualize a camera with lens cut  $\hat{\beta}$ . Direction  $\vec{\mathbf{n}}_i$ , which is the normal to the reference plane  $\Pi_i$ , is estimated by detecting line  $\mathbf{n}_i$  in every frame  $i$  and back projecting it into the 3D space.

FIG. 7 illustrates the concept of diameter  $d$ , which is the length of the segment whose endpoints are the intersections of a line that contains the principal point  $\mathbf{O}$  with the boundary. Diameter  $d$  is used to determine the Field of View (FoV)  $\Theta$  of the camera. The figure shows the particular case of the line going through the notch  $\mathbf{P}$ .

FIG. 8 schematizes the method disclosed in this disclosure for changing the direction of view by illustrating the steps that allow the rendering of the video that would be acquired by a virtual camera with predefined characteristics located in the same 3D position as a real endoscopic camera. The output of this method is a mapping function  $w$  that converts pixels  $\hat{\mathbf{u}}$  in the target image acquired by the virtual camera into pixels  $\mathbf{u}$  in the source image acquired by the real camera.

FIG. 9 illustrates how to compute point  $\bar{\mathbf{O}}_i$  in the source image, which is the point where the warping function will map the principal point of the target image  $\hat{\mathbf{O}}_i$ .  $\bar{\mathbf{O}}_i$  can be determined by transforming the origin of the reference frame of the canonical image of the target camera  $[0, 0]^T$  by a function  $g$  that is the composition of the camera model  $c_s$  and the motion  $m$  ( $g = c_s \circ m$ ). Knowing point  $\bar{\mathbf{O}}_i$  allows to estimate the limiting viewing angle that is used in the methods of FIG.

10 to prevent the existence of empty regions in the target images.

FIG. 10 shows the strategies to prevent the images rendered by the target camera from having regions without visual content (empty regions). FIG. 10 (a) shows a part of the interior of the target boundary being mapped beyond the boundary of the source image, leading to an empty region in the bottom without image content. FIG. 10 (b) shows a trivial solution that is to estimate the focal length  $\hat{f}$  such that the point that is farthest from the source boundary moves to be coincident with it. Unfortunately, this approach decreases the FoV, which is undesirable. FIG. 10 (c) shows the proposed solution where both the focal length  $\hat{f}$  and the principal point  $\hat{O}_i$  are simultaneously adjusted to remove the empty regions, while avoiding a reduction in the desired FoV.

FIG. 11 describes the different steps of the method in this disclosure for estimating the focal length and principal point of the target image that conciliate the change in the DoV with the rendering of an image with the desired FoV and no empty regions.

FIG. 12 shows how to dimension a source camera such that it properly accommodates the lens cuts and FoVs of targeted cameras by rendering realistic target images without empty regions and having the principal point close to the image center.

FIG. 13 shows the image results of rendering two images acquired with two endoscopes with different lens cuts and FoVs. The top figure shows the image acquired by the real camera, which is a laparoscope with a lens cut  $\beta = 45^\circ$  and a diameter of 10mm. The middle figure shows the virtual view obtained for  $\gamma_1 = -15^\circ$ , in which case an endoscope with lens cut  $\hat{\beta}_1 = 30^\circ$  and FoV  $\hat{\theta}_1 = 60^\circ$  is virtualized. The bottom figure depicts the virtual view obtained for  $\gamma_2 = 25^\circ$  which means that the endoscopic camera is looking downwards with respect to the symmetry axis of the scope, having  $\hat{\beta}_2 = 70^\circ$  and FoV  $\hat{\theta}_2 = 50^\circ$ .

FIG. 14 shows the image results of changing the DoV by an angle  $\gamma$  while predefining the parameters of the target camera to be the same as the calibration estimates of the source camera. The top figure is the original, real image acquired by an arthroscope with a lens cut  $\beta = 30^\circ$ . The middle figure shows the virtual view obtained for  $\gamma = -30^\circ$ , in which case the arthroscope is forward viewing along its symmetry axis ( $\beta + \gamma = 0^\circ$ ). The bottom figure depicts the virtual view obtained for  $\gamma = 35^\circ$  which means that the camera is looking downwards  $\beta + \gamma = 65^\circ$  with respect to the symmetry axis of the scope. The filled circles indicate the location of  $\hat{O}_i$ , which is the point in the source image where the principal point of the target image is mapped, showing that a change in the DoV causes a shift in the principal point.

FIG. 15 shows the image results for the directional zoom that enables to scale up the image around an arbitrary viewing direction while preserving the overall FoV and image contents. The top figure shows the original view. The middle figure shows the rendering results for  $\beta = 65^\circ$  for the case of the pre-sets of the target camera being the calibration of the source camera. The bottom figure shows the virtual view for the case of the radial distortion being  $\hat{\xi} = 5\xi$  while the lens cut is the same as in middle figure. It can be observed that by increasing the radial distortion it is possible to zoom in the region of interest while preserving all image contents that appear in the top figure. The filled circles indicate the location of  $\hat{O}_i$ , which is the point in the source image where the principal point of the target image is mapped, showing that a change in the DoV causes a shift in the principal point.

FIG. 16 is a diagrammatic view of an example computing system that includes a general purpose computing system environment.

## DETAILED DESCRIPTION

**[0023]** Systems and methods for changing the direction of view during video guided clinical procedures are disclosed herein. The systems and methods may be used for clinical procedures including, but not limited to, arthroscopy, laparoscopy, endoscopy or other surgical procedures including minimally invasive orthopedic surgery procedures. The systems and methods can be used with real-time image processing or delayed image processing.

**[0024]** FIG. 1 schematizes an endoscopic camera 34, to be used in the aforementioned video guided clinical procedures, that consists of an endoscope 10 whose proximal end 16 is assembled to a camera head 28 by means of a connector that allows the endoscope 10 to perform a rotation in azimuth by an angle  $\alpha$  18 around the mechanical axis 12. The endoscope 10 typically contains a Field Stop Mask (FSM) along the image forwarding system that causes the light that is forwarded from the distal end 14 to the proximal end 16 to form a canonical image 22 containing a circular boundary 24 and a notch 26. The camera head 28 transforms the canonical image 22 into an image in pixels 30 also presenting a circular boundary 24 and a notch 26. Moreover, and as discussed in the background section, the direction of view (DoV) 36 changes while rotating the endoscope 10 with respect to the camera head 28 because of the lens at the distal end 14 that causes the optical axis 36 and the mechanical axis 12 to be misaligned by an angle  $\beta$  denoted as lens cut 20.

**[0025]** The rotation in azimuth 18 around the mechanical axis 12 of the endoscope 10 causes the optical axis 36 to describe a cone in space 32 (the cone of DoV) whose half-angle is the lens cut 20, as illustrated in FIG. 2.

**[0026]** In this disclosure, 2D and 3D vectors are written in bold lower and upper case letters, respectively. Functions are represented by lower case italic letters, and angles by lower case Greek letters. Points and other geometric entities in the plane are represented in homogeneous coordinates, as is commonly done in projective geometry, with 2D linear

transformations in the plane being represented by 3x3 matrices and equality being up to scale. In addition, when representing functions, the symbol ; is used to distinguish between variables (that appear to the left of ;) and parameters (that appear to the right of ;) of the function. Finally, different sections of the text are referenced by their paragraphs' numbers using the symbol §.

5

## Image warping

[0027] The disclosed methods and systems for the rendering of virtual views with an arbitrary shift in the inclination of viewing direction relate with image warping techniques, in particular with software based methods to create a virtual Pan-Tilt-Zoom (PTZ) camera from a wide Field of View (FoV), panoramic camera. In this case, the image that would be acquired by the PTZ camera (the target image) is rendered from the image acquired by the panoramic camera (the source image) through a function that maps pixels in one image into pixels in the other.

10

[0028] Without loss of generality, let  $w$  be the function that transforms pixel coordinates  $u_t$  in the target image into pixel coordinates  $u_s$  in the source image, as illustrated in FIG. 3. The color value of pixel  $u_t$  in the target image can be determined by using any type of interpolation approach in the spatial or in the frequency domains, which includes, but is not limited to, nearest neighbors, next neighbors, previous neighbors, bi-linear, bi-cubic, Lanczos, bi-cubic b-spline, Mitchell-Netravali, Catmull-Rom, Kriging based, wavelet based, or edge-directed interpolation. Data driven interpolation filters learned using machine learning or deep learning can also be employed for obtaining the color value of  $u_t$ .

15

[0029] The existing warping techniques include, but are not limited to, direct mapping, inverse mapping, warping by re-sampling in the continuous or discrete image domain, warping by re-sampling and filtering, warping using a look-up table, warping using decomposable transformations and learned warping transformations.

20

[0030] The warping function  $w$  is the composition of functions  $c_s$  and  $c_t$ , corresponding to the camera models of the source and target cameras, respectively, with function  $m$ , which is the camera motion. The camera models  $c_s$  and  $c_t$  describe the mapping between the canonical image 22 in millimeters and the image in pixels 30. Since the source camera is a real camera,  $c_s$  can be determined using an appropriate calibration method. On the other hand, the target camera  $c_t$  is chosen so that the desired imaging features (resolution, zoom, FoV, etc) are predefined. Concerning function  $m$ , it describes the relative motion between virtual (target) and real (source) cameras. In more detail, it represents the rotation undergone by the virtual camera in 3D space that causes an homography mapping in projective coordinates between the canonical images of the source and target cameras.

25

[0031] Warping images acquired with endoscopic cameras is significantly more challenging than doing so with images acquired with conventional cameras because of two main reasons. Firstly, the camera model  $c_s$  changes in every frame time instant due to the relative rotation of the endoscopic lens with respect to the camera head, and this must be taken into account in building the warping function  $w$ . Secondly, the motion model  $m$  depends not only on the desired change in elevation  $\gamma$  but also in the mechanical change in azimuth  $\delta$  that must be measured at every frame time instant. These challenges do not exist for conventional cameras because they do not present moving parts that interfere with the camera model.

30

35

## Endoscopic camera model

[0032] This section introduces the endoscopic camera model  $c$  by describing the model of a general camera presenting radial distortion introduced by the optics, providing an overview of relevant concepts and explaining how the endoscopic camera can be described with an adaptive model that is updated at every frame time instant.

40

### General camera model

45

[0033] FIG. 4 schematizes the different steps that comprise the camera model function  $c$ . A 3D point  $X = [X \ Y \ Z]^T$  is projected onto the image plane according to the pinhole model, yielding a 2D point  $x$  in the canonical image 22. This projection is denoted as  $p$  and corresponds to computing  $x = p(X) = [X/Z, Y/Z]^T$ . Afterwards, distortion  $d$  is an operation performed in the canonical image 22 that applies distortion to point  $x$ , yielding point  $x'$ :  $x' = d(x; \xi)$ , with  $\xi$  being the distortion parameters. As a final step, the application of the intrinsic parameters converts points in millimeters  $x'$  into points  $u$  in the image in pixels 30. This is accomplished through function  $k$  by computing  $u = k(x'; f, \mathbf{O}) = \mathbf{f}x' + \mathbf{O}$ , with  $f$  being the focal distance,  $\mathbf{O}$  the principal point and assuming, without loss of generality, that the pixels are squared, i.e., the aspect ratio is unitary and the skew is zero. In case this assumption is not made, it is straightforward to update function  $k$  to include two extra parameters that account for the aspect ratio and the skew. The camera model  $c$  is a function applied in 2D space that transforms points  $x$  in the canonical image 22 into points  $u$  in the image in pixels 30,  $u = c(x; f, \mathbf{O}, \xi)$ , and is obtained by the composition of the distortion  $d$  and the intrinsics  $k$ :  $c = k \circ d$ .

50

55

[0034] In this model, the distortion function  $d$  is generic, meaning that it can be any distortion model in the literature such as Brown's polynomial model, the rational model, the fish-eye model, or the division model with one or more parameters, in

which case  $\xi$  is a scalar or a vector, respectively. The distortion function  $d$  can also take into account other types of distortion such as, but not limited to, tangential distortion, prism distortion or perspective distortion caused by tilt of the sensor. Without loss of generality, it is assumed in the remainder of the description that  $d$  applies the first order division model, with

$$\mathbf{x}' = d(\mathbf{x}; \xi) = \begin{bmatrix} x \\ y \end{bmatrix} \frac{1}{1 + \xi \sqrt{x^2 + y^2}}.$$

$\xi$  being a scalar:

**[0035]** Moreover, it is known that a distance  $r$  in pixels measured in the image in pixels 30 between a point  $\mathbf{u}$  and the principal point  $\mathbf{O}$  ( $r = \|\mathbf{u} - \mathbf{O}\|$ ) corresponds to an angle  $\theta$  in 3D that depends on the joint effect of  $f$  and  $\xi$ , such that

$$\Phi(f, \xi, \theta, r) = 0. \quad (1)$$

#### Handling lens rotation

**[0036]** The rotation of the endoscope 10 with respect to the camera head 28 causes the principal point  $\mathbf{O}$ , the center  $\mathbf{C}$  of the circular boundary 24 and the notch 26, denoted as  $\mathbf{P}$ , to rotate in the image in pixels 30 around a point  $\mathbf{Q}$ , as illustrated in FIG. 5. In fact,  $\mathbf{O}$  and  $\mathbf{C}$  rotate around  $\mathbf{Q}$  because they are usually not coincident due to mechanical tolerances during lens manufacturing. Knowing that  $\mathbf{O}$  is the image of the optical axis,  $\mathbf{C}$  is the center of the circular boundary that depends on how the FSM is placed with respect to the lens rod and  $\mathbf{Q}$  is the intersection of the mechanical axis 12 with the image plane, which depends on the eye-piece, the coincidence of all three points occurs when the eye-piece and the FSM are perfectly aligned with the lens rod. The misalignment between  $\mathbf{Q}$  and  $\mathbf{C}$  causes the boundary to change positions in the image and  $\mathbf{C}$  to move in a circle around  $\mathbf{Q}$  while the lens rotates. When  $\mathbf{Q}$  and  $\mathbf{O}$  are misaligned,  $\mathbf{O}$  also describes a circular trajectory around  $\mathbf{Q}$ .

**[0037]** Thus, if the calibration parameters  $f$ ,  $\mathbf{O}$ ,  $\xi$  refer to a particular azimuth position  $\alpha_0$ , then a rotation in azimuth by an angle  $\delta_i = \alpha_i - \alpha_0$  changes the position of the principal point to  $\mathbf{O}_i = r(\mathbf{O}; \delta_i, \mathbf{Q})$ , where  $r$  denotes a 2D rotation by an angle  $\delta_i$  around an axis going through  $\mathbf{Q}$ . Similarly, the center of the circular boundary becomes  $\mathbf{C}_i = r(\mathbf{C}; \delta_i, \mathbf{Q})$ .

**[0038]** In other words, let  $f$ ,  $\mathbf{O}$  and  $\xi$  be the calibration parameters of a generic endoscopic camera at a particular azimuth position  $\alpha_0$ . It comes that the camera model that maps points  $\mathbf{x}$  in the canonical image into points  $\mathbf{u}$  in pixel coordinates can change at every frame time instant  $i$ , being given by  $\mathbf{u} = c(\mathbf{x}; f, \mathbf{O}_i, \xi)$ , with  $\mathbf{O}_i = r(\mathbf{O}; \delta_i, \mathbf{Q})$  and  $\delta_i = \alpha_i - \alpha_0$ , where  $\alpha_i$  is the angle in azimuth in frame  $i$ .

**[0039]** In the remainder of this description, it is assumed, without loss of generality, that both the calibration parameters  $f$ ,  $\mathbf{O}$ ,  $\xi$  for a reference azimuth position  $\alpha_0$  and the rotation center  $\mathbf{Q}$  are known *a priori*, such that if rotation  $\delta_i$  with respect to  $\alpha_0$  is estimated at each frame time instant  $i$ , then the calibration parameters  $f$ ,  $\mathbf{O}_i$ ,  $\xi$  can be determined as previously described. It is relevant to note that if  $\mathbf{O}$  is coincident with  $\mathbf{Q}$ , then  $\mathbf{O}_i$  in every frame  $i$  will also be coincident with  $\mathbf{Q}$  and thus this adaptation of the camera model  $c$  to the rotation of the scope is unnecessary. However, the misalignment between these entities occurs frequently due to mechanical tolerances in building the optics, as discussed previously. Since the effect of this misalignment in the calibration parameters is not negligible, it must be taken into account in practice.

#### Camera rotation motion in 3D

**[0040]** The method to determine the motion model  $m$  introduced in §§ [0043]-[0048] is now described, with the accompanying FIG. 6 providing an illustrative scheme to better explain the most relevant concepts. As previously mentioned, a change in the DoV of a generic camera can be accomplished by rotating the camera around an axis going through its projection center. In our particular case of endoscopic cameras, the virtualization of a camera with lens cut  $\beta$  from a real camera with lens cut  $\beta$  corresponds to a change in the DoV that can be accomplished by virtually rotating the real camera in 3D about its center of projection by an angle  $\gamma = \beta - \beta$  around a direction  $\vec{\mathbf{n}}_i$  42. This direction  $\vec{\mathbf{n}}_i$  42 is the normal to a plane  $\Pi_i$  40, henceforth referred to as the *reference plane*, that is defined by the mechanical and optical axes, which, in ideal conditions, are coplanar but not coincident because of the non-zero lens cut. Since the optical axis 36 defines the DoV that changes with the rotation in azimuth 18, the motion model  $m$ , which is a 3D rotation of  $\gamma$  around axis  $\vec{\mathbf{n}}_i$  42, can vary at every frame time instant  $i$  and must be estimated accordingly.

**[0041]** As illustrated in FIG. 6, the reference plane  $\Pi_i$  40 is projected into a line  $\mathbf{n}_i$  38 in the image in pixels 30. Thus, it is possible to determine  $\vec{\mathbf{n}}_i$  on-the-fly by detecting line  $\mathbf{n}_i$  38 in every frame  $i$  and back-projecting it into the 3D space.

**[0042]** In ideal conditions, plane  $\Pi_i$  40 contains both the optical and the mechanical axes and intersects the FSM in the notch, whose purpose is to inform the surgeon about the direction of the lens cut. These conditions hold if and only if points  $\mathbf{Q}$ ,  $\mathbf{O}_i$  and  $\mathbf{P}_i$  in the image in pixels 30 are collinear and, in this case, line  $\mathbf{n}_i$  38 is the line that contains all three points  $\mathbf{Q}$ ,  $\mathbf{O}_i$  and  $\mathbf{P}_i$ . In real conditions, this does not usually happen due to the mechanical tolerances in lens manufacturing, leading to a non-coplanarity of the optical and mechanical axes and/or plane  $\Pi_i$  40 not going exactly through the notch of the FSM. Thus, in the remainder of this description, it is assumed, without loss of generality, that  $\mathbf{n}_i$  38 is the line defined by the optical

axis  $\mathbf{O}_i$  and the notch  $\mathbf{P}_i$ . Thus, in projective coordinates, line  $\mathbf{n}_i$  38 is computed by  $\mathbf{n}_i = \mathbf{P}_i \times \mathbf{O}_i$  and the normal to plane  $\Pi_i$  40 can afterwards be determined in a straightforward manner by  $\vec{\mathbf{n}}_i = \mathbf{K}_i^T \mathbf{n}_i$ , with  $\mathbf{K}_i$  being the matrix of intrinsic

5

$$\mathbf{K}_i = \begin{bmatrix} f & 0 & O_{ui} \\ 0 & f & O_{vi} \\ 0 & 0 & 1 \end{bmatrix}$$

parameters in frame  $i$ , and  $\mathbf{O}_i = [O_{ui}, O_{vi}]^T$ .

10

**[0043]** Alternatively, since in ideal conditions points  $\mathbf{Q}$ ,  $\mathbf{O}_i$  and  $\mathbf{P}_i$  in the image in pixels 30 are collinear, as is the boundary center  $\mathbf{C}_i$ , line  $\mathbf{n}_i$  38 can be obtained by determining the line that goes through points  $\mathbf{Q}$  and  $\mathbf{O}_i$  or  $\mathbf{C}_i$  and  $\mathbf{O}_i$ , in which case points  $\mathbf{Q}$  and  $\mathbf{C}_i$  replace  $\mathbf{P}_i$  in the previous equation, respectively.

**Further considerations**

15

**[0044]** In this disclosure, it is assumed, without loss of generality, that all measurements and computations are performed in dry environment. Adaptation to wet environment can be performed in a straightforward manner by multiplying the focal length by the ratio of the refractive indices of the two or more mediums where light travels before reaching the imaging sensor.

20

**[0045]** Another important consideration to make is that both the update of the camera model  $c$  and the detection of line  $\mathbf{n}_i$  38 in the image in pixels 30 in every frame  $i$  can be performed by determining the angular displacement in azimuth  $\delta_i$  with respect to a reference angular position  $\alpha_0$ , which may be accomplished using exclusively image processing techniques. In this disclosure is it considered, without loss of generality, that the method disclosed in US Application No. 62/911,950 for detecting the boundary with center  $\mathbf{C}_i$  and notch  $\mathbf{P}_i$  in every frame  $i$  is employed for this task. This method contemplates the possibility that the FSM contains more than one notch, which is useful for guaranteeing that at least one notch is always visible in the image, being able to detect multiple notches whose relative location is known and that are identified by their different shapes and sizes. Thus, it comes that the angular displacement  $\delta_i$  is the angle defined by points  $\mathbf{P}_i$ ,  $\mathbf{Q}$  and  $\mathbf{P}$  ( $\delta_i = \sphericalangle \mathbf{PQP}_i$ ), with  $\mathbf{P}$  being the position of one of the notches at the reference angular position  $\alpha_0$ . Alternatively,  $\delta_i$  can be estimated from the boundary centers  $\mathbf{C}_i$  and  $\mathbf{C}$ , with  $\mathbf{C}$  being the center of the boundary in the reference position:  $\delta_i = \sphericalangle \mathbf{CQC}_i$ .

30

**[0046]** As discussed in § [0054], a distance  $r$  measured in the image in pixels 30 between any point and the principal point  $\mathbf{O}$  corresponds to an angle  $\theta$  in 3D that depends on both  $f$  and  $\xi$  and that can be computed using Equation 1. Let  $d$  be the length of the line segment whose endpoints are the two intersections of the boundary with the line  $\mathbf{n}$  38 in FIG. 7 that goes through the principal point  $\mathbf{O}$ . This length  $d$  is referred to as *diameter* despite not corresponding to the diameter of the circular boundary, as this only happens in case the principal point  $\mathbf{O}$  is coincident with the boundary center  $\mathbf{C}$ . As illustrated in FIG. 7,  $d$  is the addition of a lower and an upper radius:  $d=r_l+r_u$ . Using both  $r_l$  and  $r_u$ , angles  $\theta_l$  and  $\theta_u$  can be determined from Equation 1. The field of view (FoV) of the camera  $\Theta$  is defined as the addition of these two angles:  $\Theta = \theta_l + \theta_u$ . Any line  $\mathbf{n}$  going through the principal point  $\mathbf{O}$  defines a different diameter  $d$ , and FIG. 7 shows the particular case of  $\mathbf{n}$  going through the notch  $\mathbf{P}$ .

35

**[0047] Method for changing the Direction of View (Do V)**

40

**[0048]** This section presents the method disclosed in this disclosure for changing the DoV by rendering the video that would be acquired by a virtual camera with predefined characteristics located in the same 3D position as a real endoscopic camera. In particular, let  $I_i$  be frame  $i$  acquired by the real endoscopic camera, henceforth referred to as the *source camera*, whose endoscope has a lens cut  $\beta$  and rotates in azimuth around a mechanical axis that intersects the image plane in the rotation center  $\mathbf{Q}$ . The endoscopic camera's calibration for a reference angular position  $\alpha_0$ , corresponding to a particular notch position  $\mathbf{P}$ , is known, meaning that its focal length  $f$ , radial distortion  $\xi$  and principal point  $\mathbf{O}$  have been determined. The purpose of this method is to render an image  $\hat{I}_i$  with resolution  $m \times n$  and a circular boundary with diameter  $\hat{d}$  centered in point  $\hat{\mathbf{C}}$  that would be acquired by a virtual camera, henceforth referred to as the *target camera*, with a lens cut  $\hat{\beta}$ , Field-of-View  $\hat{\Theta}$ , and distortion  $\hat{\xi}$  placed in the same 3D location as the source camera.

45

**[0049]** As shown in FIG. 8, the method starts by detecting the boundary with center  $\mathbf{C}_i$  and notch  $\mathbf{P}_i$  in image  $I_i$ . By determining the rotation in azimuth  $\delta_i$  as the angle defined by points  $\mathbf{P}_i$ ,  $\mathbf{Q}$  and  $\mathbf{P}$ , it is possible to estimate the location of the principal point  $\mathbf{O}_i$  by rotating  $\mathbf{O}$  by an angle  $\delta_i$  around  $\mathbf{Q}$ , as described in §§ [0055]-[0059]. This provides the updated model for the source camera  $c_s$  that maps points  $\mathbf{x}$  in the canonical image, which are the pinhole projection of points  $\mathbf{X}$  in the 3D scene, into points  $\mathbf{u}$  in the image in pixels, such that  $\mathbf{u}=c_s(\mathbf{x}; f, \mathbf{O}_i, \xi)$ .

50

55

**[0050]** Line  $\mathbf{n}_i$ , which is the projection of the reference plane  $\Pi_i$  introduced in §§ [0060]-[0064] in the image plane, is determined as the line that goes through points  $\mathbf{O}_i$  and  $\mathbf{P}_i$ . As previously described, by back-projecting  $\mathbf{n}_i$  into the 3D space, plane  $\Pi_i$  with normal  $\vec{\mathbf{n}}_i$  can be obtained, yielding motion model  $m$ , which is a rotation in 3D space around axis  $\vec{\mathbf{n}}_i$  by an angle  $\gamma$  that corresponds to the difference between the lens cuts of the target and source cameras,  $\gamma = \hat{\beta} - \beta$ . Motion model  $m$  transforms points  $\hat{\mathbf{x}}$  into points  $\mathbf{x}$  in the canonical images of the target and source cameras, respectively, such that  $\mathbf{x} = m\hat{\mathbf{x}}$ ;

$\gamma, \vec{n}_i$ ).

[0051] In order to derive the camera model for the target camera  $ct$  that maps points  $\hat{x}$  in its canonical image into points  $\hat{u}$  in the image in pixels ( $\hat{u} = c_t(x; \hat{f}, \hat{O}_i, \xi)$ ), the focal length  $\hat{f}$  and the location of the principal point  $\hat{O}_i$  in every frame  $i$  must be computed. Details on the estimation of these parameters are given in §§ [0079]-[0091].

[0052] As a final step, the algorithm generates the target image  $\hat{I}_i$  using image warping techniques, as described in §§ [0043]-[0048], where each pixel  $\hat{u}$  in  $\hat{I}_i$  is mapped into a point  $u$  in the source image  $I_i$  by the mapping function  $w$ , that is the

composition of functions  $c_s$ ,  $m$  and  $c_t^{-1}$  ( $w = c_s \circ m \circ c_t^{-1}$ ), such that the color value of  $w$  can be interpolated. As previously described, this mapping function  $w$  may implement any method for image warping or pixel value interpolation.

[0053] The disclosed method for changing the DoV assumes that the calibration parameters of the source camera for the reference position in azimuth  $\alpha_0$  are known in advance. These can be determined in several different ways, which include, but are not limited to, using a set of calibration parameters predetermined in factory or representative of a set of similar endoscopic cameras and using an appropriate calibration method for performing calibration in the Operating Room before the medical procedure, such as the one disclosed in U.S. Patent No. 9,438,897 (application no. 14/234,907) entitled "Method and apparatus for automatic camera calibration using one or more images of a checkerboard pattern".

[0054] In addition, the presently disclosed method also assumes that the rotation center  $Q$  is known a priori. However, this is not a strict requirement since  $Q$  may be determined on-the-fly from points  $P_i$  and/or  $C_i$  detected in successive frames, using the method disclosed in US Application No. US Application No. 62/911,950. Another alternative is to determine the position of the principal point  $O_i$  in every frame time instant  $i$  by making use of an estimate for the principal point given in a normalized reference frame that is attached to the circular boundary. In this case, since  $O_i$  is obtained directly from the normalized estimate of the principal point, as disclosed in US Application No. US Application No. 62/911,950, it is not required to know the rotation center  $Q$  or the angular displacement in azimuth  $\delta_i$  a priori.

[0055] The estimation of the angular displacement in azimuth  $\delta_i$ , depicted in the second block of the diagram in FIG. 8, is performed using exclusively image processing methods. However, there are alternative estimation approaches that can be employed to measure the changes in rotation in azimuth of the scope with respect to the camera-head at every frame time instant. These include the use of sensing devices such as a rotary encoder attached to the camera head or an optical tracking system for determining the position of an optical marker attached to the scope cylinder.

[0056] The method for changing the DoV disclosed herein may be used for different purposes by setting the parameters of the target camera to desired values. In particular, if the distortion of the target camera  $\xi$  is set to zero, the images  $\hat{I}_i$  rendered by the target camera will be distortion-free images. Also, in order to perfectly mimic a real endoscopic camera comprising an FSM, the rendered target images  $\hat{I}_i$  may be processed to create a black frame such that meaningful image contents are within a circular region with center  $\hat{C}$  and diameter  $\hat{d}$  and to create a notch by placing a visual mark in point

$\hat{P}_i = \hat{C}_i + \hat{d}/2 \mathbf{v}_i$  in the circular boundary, with  $\mathbf{v}_i$  being the unit direction of line  $\mathbf{n}_i$  ( $\mathbf{v}_i = \frac{\mathbf{P}_i - \mathbf{O}_i}{\|\mathbf{P}_i - \mathbf{O}_i\|}$ ), computed in block 4 of the diagram in FIG. 8. The image boundary can also take any other desired geometric shape, such as, but not limited to, a conic shape, a rectangular shape, a hexagonal shape or any other polygonal shape.

### Adjusting the target camera model

[0057] The disclosed warping function  $w$  is the composition of three functions: the camera model  $c_s$ , for which the parameters are the calibration parameters of the real source camera at the current frame  $i$ , the motion model  $m$  that depends on the desired change in inclination  $\gamma$  and orientation  $\vec{n}_i$  of the reference plane with respect to camera-head, and the camera model  $ct$  for the virtual target camera that must be such that the FoV is  $\hat{\Theta}$  for an image distortion of  $\xi$  and image diameter  $\hat{d}$ .

[0058] While the two first functions are fully determined, it remains to define the focal length  $\hat{f}$  and principal point  $\hat{O}_i$  of the target camera to fully specify  $c_t$ .

[0059] Since both functions  $c_s$  and  $m$  are known, it is possible to anticipate the location  $\bar{O}_i$  where the principal point of the target image will be mapped by the warp function  $w$  in the source image (FIG. 9). Considering  $g$  to be the composition of  $c_s$  and  $m$ , such that  $g = c_s \circ m$ , it can be shown that the location is  $\bar{O}_i = g(\mathbf{0})$  with  $\mathbf{0}$  being the origin  $[0, 0]^T$ . This location is independent of the values of  $\hat{f}$  and  $\hat{O}_i$  that have not been selected yet.

[0060] The choice of parameters  $\hat{f}$  and  $\hat{O}_i$  for the target camera  $c_t$  can be performed according to three different settings, for which an illustration is given in FIG. 10, as follows.

[0061] *Setting A:* One possibility is to choose  $\hat{O}_i$  coincident with the center  $\hat{C}$  of the boundary. Since the virtual center of rotation  $\hat{Q}$  is also assumed to be coincident with  $\hat{C}$  then the principal point is stationary across frames  $i$ , independently of the rotation of the lens with respect to camera-head. In this case, and considering the interdependence between focal length, distortion, image diameter and FoV expressed by  $\Phi$ , the focal length  $\hat{f}$  that grants the specified FoV of  $\hat{\Theta}$  is the solution of  $\Phi(\hat{f}, \hat{\xi}, \hat{\Theta}/2, \hat{d}/2) = 0$ .

[0062] The problem of this choice of  $\hat{f}$  and  $\hat{\mathbf{O}}_i$  is that the rendered target image often has a region without visual content (the *empty region*) near the intersection of  $\mathbf{n}_i$  with the boundary (FIG. 10 (a) RIGHT). This happens whenever angle  $\bar{\Psi}_i$ , the *limit viewing angle*, defined by the back-projection rays of  $\mathbf{O}_i$  and its closest point in the boundary of the source image  $\hat{\mathbf{S}}_i$ , is smaller than half of the desired FoV  $\hat{\Theta}$ .

[0063] The empty region arises whenever the selected warping function  $w$  maps the target image beyond the boundary of the source image where there are no visual contents (FIG. 10(a) LEFT). This is obviously an undesirable artifact that undermines the user experience in varying the DoV via software.

[0064] *Setting B*: A possible solution to avoid the empty region artifact is to relax the FoV requirement and make it twice the limiting viewing angle  $\bar{\Psi}_i$  (FIG. 10 (b) RIGHT). In this case, the principal point is still kept in the center of the boundary ( $\hat{\mathbf{O}}_i = \hat{\mathbf{C}}$ ) and  $\hat{f}$  is determined by solving  $\Phi(\hat{f}, \hat{\xi}, \bar{\Psi}_i, \hat{d}/2) = 0$  with respect to  $\hat{f}$ . Setting B has the drawback that the FoV of the target image can be substantially smaller than what was specified, which is in general an issue in terms of application (FIG. 10(b) LEFT).

[0065] *Setting C*: This disclosure discloses an alternative method to choose the focal length and principal point of the target image that conciliates the change in inclination of the DoV by an angle  $\gamma$ , with the rendering of an image with the desired FoV  $\hat{\Theta}$  and no empty regions (FIG. 10(c) RIGHT). The different steps that comprise this novel method are provided in the scheme of FIG. 11. The strategy serves to compensate the lack of visual contents in the top or bottom of the source image by the visual contents that are available below or above, in case  $\gamma < 0$  or  $\gamma > 0$ , respectively.

[0066] For illustrative purposes let the shift in the DoV be a positive angular offset  $\gamma > 0$  (FIG. 10(c) LEFT) such that the limiting viewing angle  $\bar{\Psi}_i$  is measured towards the notch (the down direction). Since the FoV of the target image in the down direction has to be decreased by an amount  $\Theta - \bar{\Psi}_i$  to avoid the creation of an empty region, the idea is to increase the FoV in the up direction by the same amount to obtain an overall FoV of  $\hat{\Theta}$  as in the specified requirements for the target camera. The shift of FoV from down to up directions causes the principal point to shift downwards by an amount  $\lambda$ , such that  $\hat{\mathbf{O}}_i = \hat{\mathbf{C}} + \lambda \mathbf{v}_i$ . In this case, the focal length  $\hat{f}$  and  $\lambda$  can be determined by solving the following system of equations

$$\begin{cases} \Phi(\hat{f}, \hat{\xi}, \bar{\Psi}_i, \hat{d}/2 - \lambda) = 0 \\ \Phi(\hat{f}, \hat{\xi}, \hat{\Theta} - \bar{\Psi}_i, \hat{d}/2 + \lambda) = 0 \end{cases}$$

[0067] Note that if  $\gamma < 0$ , the shift in the FoV is from up to down directions and the principal point translates upwards, i.e.,  $\hat{\mathbf{O}}_i = \hat{\mathbf{C}} - \lambda \mathbf{v}_i$ .

[0068] If  $\Theta \geq 0$ , with  $\Theta$  being the FoV of the source camera, then the disclosed method for selecting  $\hat{f}$  and  $\hat{\mathbf{O}}_i$  always succeeds in rendering a target image with the desired FoV and no empty region. Conciliating these two features is important to provide a good user experience, but there are situations in which this is accomplished by significantly deviating the principal point towards the periphery of the target image. This might be undesirable for certain applications, specially the ones where the target camera is intended to mimic a particular real endoscopic camera, in which case the principal point is typically close to the image the center.

### Applications and functionalities

[0069] The disclosed image processing methods for changing the DoV of endoscopic systems with exchangeable, rotatable optics can lead to multiple applications and be employed in several different systems. The disclosure describes some embodiments of these systems and applications, without prejudice of other possible applications.

#### *Electronic switch between rigid endoscopes with different lens cut $\beta$*

[0070] Manufacturers of rigid endoscopes provide lenses with different angular offsets from the mechanical axis with the most common lens cuts being  $\beta = 0^\circ, 30^\circ, 45^\circ, 70^\circ$ . Although surgical procedures usually favor endoscopes with a particular cut angle, there are moments or steps in the surgery where a different lens cut would be more convenient. For example, in knee arthroscopy the preference goes for arthroscopes with  $30^\circ$  lens cut, but the meniscus inspection in anterior or intercondylar regions benefits from a cut angle of  $0^\circ$  or  $70^\circ$ .

[0071] The problem is that, since a lateral change in DoV requires to physically switch the endoscope, which causes disruption and can involve risks for the patient, the surgeons rarely do it in practice and perform the procedure with the same endoscope, even when the visualization is sub-optimal.

[0072] The disclosed method can be used in a system to process the images and video acquired by an endoscopy camera equipped with a lens with a wide FoV to empower the surgeon with electronic switch between two or more virtual endoscopes with different lens cuts  $\beta$ . Such system overcomes the above-mentioned difficulty, that precludes the surgeon from having the best possible visualization at every surgical moment or step.

[0073] Let the two desired virtual endoscopes have lens cuts  $\hat{\beta}_1$  and  $\hat{\beta}_2$  and FoVs  $\hat{\Theta}_1$  and  $\hat{\Theta}_2$ , respectively. Although the

disclosed method for selecting the focal length and principal point (FIG. 10 (c) and FIG. 11) delivers the desired FoV for most settings of source camera, it might happen that the principal point deviates too much from the center of the target image, which diminishes the user experience in terms of realistic virtualization of endoscopes with different lens cuts. The strategy to prevent this issue is to have a source camera with a lens cut  $\beta$  and FoV  $\Theta$  that can properly accommodate the lens cuts and FoVs of the targeted cameras. As shown in FIG. 12, and given the specifications  $\hat{\beta}_1, \hat{\Theta}_1$  and  $\hat{\beta}_2, \hat{\Theta}_2$  of the two

target cameras, this is accomplished if  $\beta$  and  $\Theta$  are approximately 
$$\beta = \frac{\hat{\beta}_1 + \hat{\beta}_2}{2} + \frac{\hat{\Theta}_1 - \hat{\Theta}_2}{4} \quad \text{and}$$

$$\Theta = \frac{\hat{\Theta}_1 + \hat{\Theta}_2}{2} + \hat{\beta}_2 - \hat{\beta}_1.$$

**[0074]** An embodiment is to use the disclosed methods in a system to electronically switch between the two most common cut angles used in arthroscopy. Since a standard 30° arthroscope has  $\hat{\beta}_1 = 30^\circ$ ,  $\hat{\Theta}_1 = 110^\circ$  and a 70° arthroscope has  $\hat{\beta}_2 = 70^\circ$ ,  $\hat{\Theta}_2 = 90^\circ$ , then it stems from the calculations above that the ideal source camera should have a cut angle  $\beta$  around 45° and FoV  $\Theta$  of approximately 140°. These values are merely indicative and serve to assure the rendering of realistic target images with the principal point in the center region. Let  $f, \xi, \mathbf{O}_i$  be the exact calibration of the source camera, and let  $\hat{d}_1, \hat{d}_2$  and  $\hat{\xi}_1, \hat{\xi}_2$  be, respectively, the image diameters and distortions that complement the specifications for the two target cameras, and that can be either arbitrary or obtained from the calibration of real lenses. The method of FIG. 8 is used to process the images acquired by the real source camera where a user command switches the settings of the target camera to alternate between the 30° lens (target 1) and the 70° lens (target 2) (FIG. 13).

#### Electronic Change of Do V, Distortion Correction and Directional Zoom

**[0075]** So far the method of FIG. 8 has been used to select a part of the FoV of the source camera and mimic the video output that would be acquired by a camera with smaller FoV and different lens cut placed in the same 3D location.

**[0076]** Another possible embodiment or application is to use the disclosed method to shift the DoV of a particular endoscopic camera while maintaining the overall FoV, in which case the principal point in the rendered images moves towards the periphery as the shift in DoV increases (FIG. 14). In this case the disclosed methods of FIG. 8 and FIG. 11 are executed in a system that is connected to the source camera, where the distortion  $\xi$ , image diameter  $\hat{d}$ , and FoV  $\hat{\Theta}$  of the target camera are set with the same values as the source camera, and where the angular shift  $\gamma$  in the DoV is commanded by the user that can freely move it both upwards or downwards (FIG. 14).

**[0077]** The system that is described can be further enhanced with the possibility of user controlled distortion, in which case the setting  $\hat{\xi} = 0, \gamma = 0$  will cause the system to correct the radial distortion in the source video.

**[0078]** The disclosed system can also be used to implement what is referred to as *directional zoom*, in which case the angular shift  $\gamma$  is adjusted such that the principal point in the rendered image becomes overlaid with a region of interest (ROI) and the distortion  $\hat{\xi}$  is increased to magnify the ROI while maintaining the FoV and all visual contents in the image (FIG. 15). As an alternative, the magnification of the ROI can also be accomplished by maintaining the distortion and increasing the focal length  $\hat{f}$  in which case the FoV decreases and some image contents will be lost.

#### Other applications

**[0079]** The disclosed methods of the schemes of FIG. 8 and FIG. 11 can be used to process a discrete set of images for the purpose of rendering images of virtual target cameras with different, known calibrations and/or lens cut angles to be used as input to train and/or test machine learning or deep learning based algorithms for automatic learning to identify and estimate different camera and/or lens characteristics, including, but not limited to, identification of the lens cut, estimation of the calibration parameters, learning to generate virtual images with different lens cuts.

**[0080]** The disclosed methods can also be applied to other types of imagery such as fundus images in ophthalmology.

**[0081]** The method that is disclosed for selecting the focal length and position of principal point in the target image (FIG. 11) is not limited to endoscopy applications and can be employed in the warping of generic images and video. One possible embodiment is the use of the method in systems of visual surveillance that implement electronic pan-and-tilt in images and videos acquired by a fish-eye camera.

**[0082]** FIG. 16 is a diagrammatic view of an illustrative computing system that includes a general purpose computing system environment 1200, such as a desktop computer, laptop, smartphone, tablet, or any other such device having the ability to execute instructions, such as those stored within a non-transient, computer-readable medium. Furthermore, while described and illustrated in the context of a single computing system 1200, those skilled in the art will also appreciate that the various tasks described hereinafter may be practiced in a distributed environment having multiple computing systems 1200 linked via a local or wide-area network in which the executable instructions may be associated with and/or executed by one or more of multiple computing systems 1200. Computing system environment 1200, or portions thereof, may find use for the processing, methods, and computing steps of this disclosure.

**[0083]** In its most basic configuration, computing system environment 1200 typically includes at least one processing unit 1202 and at least one memory 1204, which may be linked via a bus 1206. Depending on the exact configuration and type of computing system environment, memory 1204 may be volatile (such as RAM 1210), non-volatile (such as ROM 1208, flash memory, etc.) or some combination of the two. Computing system environment 1200 may have additional features and/or functionality. For example, computing system environment 1200 may also include additional storage (removable and/or non-removable) including, but not limited to, magnetic or optical disks, tape drives and/or flash drives. Such additional memory devices may be made accessible to the computing system environment 1200 by means of, for example, a hard disk drive interface 1212, a magnetic disk drive interface 1214, and/or an optical disk drive interface 1216. As will be understood, these devices, which would be linked to the system bus 1206, respectively, allow for reading from and writing to a hard disk 1218, reading from or writing to a removable magnetic disk 1220, and/or for reading from or writing to a removable optical disk 1222, such as a CD/DVD ROM or other optical media. The drive interfaces and their associated computer-readable media allow for the nonvolatile storage of computer readable instructions, data structures, program modules and other data for the computing system environment 1200. Those skilled in the art will further appreciate that other types of computer readable media that can store data may be used for this same purpose. Examples of such media devices include, but are not limited to, magnetic cassettes, flash memory cards, digital videodisks, Bernoulli cartridges, random access memories, nano-drives, memory sticks, other read/write and/or read-only memories and/or any other method or technology for storage of information such as computer readable instructions, data structures, program modules or other data. Any such computer storage media may be part of computing system environment 1200.

**[0084]** A number of program modules may be stored in one or more of the memory/media devices. For example, a basic input/output system (BIOS) 1224, containing the basic routines that help to transfer information between elements within the computing system environment 1200, such as during start-up, may be stored in ROM 1208. Similarly, RAM 1210, hard drive 1218, and/or peripheral memory devices may be used to store computer executable instructions comprising an operating system 1226, one or more applications programs 1228 (such as an application that performs the methods and processes of this disclosure), other program modules 1230, and/or program data 1232. Still further, computer-executable instructions may be downloaded to the computing environment 1200 as needed, for example, via a network connection.

**[0085]** An end-user, e.g., a customer, retail associate, and the like, may enter commands and information into the computing system environment 1200 through input devices such as a keyboard 1234 and/or a pointing device 1236. While not illustrated, other input devices may include a microphone, a joystick, a game pad, a scanner, etc. These and other input devices would typically be connected to the processing unit 1202 by means of a peripheral interface 1238 which, in turn, would be coupled to bus 1206. Input devices may be directly or indirectly connected to processor 1202 via interfaces such as, for example, a parallel port, game port, firewire, or a universal serial bus (USB). To view information from the computing system environment 1200, a monitor 1240 or other type of display device may also be connected to bus 1206 via an interface, such as via video adapter 1242. In addition to the monitor 1240, the computing system environment 1200 may also include other peripheral output devices, not shown, such as speakers and printers.

**[0086]** The computing system environment 1200 may also utilize logical connections to one or more computing system environments. Communications between the computing system environment 1200 and the remote computing system environment may be exchanged via a further processing device, such a network router 1252, that is responsible for network routing. Communications with the network router 1252 may be performed via a network interface component 1254. Thus, within such a networked environment, e.g., the Internet, World Wide Web, LAN, or other like type of wired or wireless network, it will be appreciated that program modules depicted relative to the computing system environment 1200, or portions thereof, may be stored in the memory storage device(s) of the computing system environment 1200.

**[0087]** The computing system environment 1200 may also include localization hardware 1256 for determining a location of the computing system environment 1200. In embodiments, the localization hardware 1256 may include, for example only, a GPS antenna, an RFID chip or reader, a Wi-Fi antenna, or other computing hardware that may be used to capture or transmit signals that may be used to determine the location of the computing system environment 1200.

**[0088]** In a first aspect of the instant disclosure, a method is disclosed that, given video acquired by a real source camera that comprises a camera-head and a rigid endoscope with lens cut  $\beta$  that rotates in azimuth around a mechanical axis that intersects the image in a point  $\mathbf{Q}$ , for which the focal length  $f$ , radial distortion  $\xi$ , and principal point  $\mathbf{O}$  at a certain azimuth  $\alpha_0$  are known, renders the video that would be acquired by a virtual target camera placed in the same 3D location as the source camera, but with a lens cut  $\hat{\beta}$ , Field-of-View  $\hat{\theta}$ , and distortion  $\xi$ , where each successive target image  $\hat{I}_i$  with resolution of  $m \times n$  and circular boundary of diameter  $\hat{d}$  centered in point  $\hat{\mathbf{C}}$  is rendered by executing the following steps for each corresponding source image  $I_i$ : (i) finding the location of the principal point  $\mathbf{O}_i$  in the source image  $I_i$  with azimuth  $\alpha_i$  by rotating  $\mathbf{O}$  around  $\mathbf{Q}$  by an angular displacement in azimuth  $\delta_i = \alpha_i - \alpha_0$ ; (ii) updating the camera model of the source camera that maps points  $\mathbf{x}$  in the canonical image, which are the pinhole projection of points  $\mathbf{X}$  in the 3D scene, into points  $\mathbf{u}$  in the pixel image such that  $\mathbf{u} = c_s(\mathbf{x}; f, \mathbf{O}_i, \xi)$ ; (iii) determining the 3D location of a vertical plane  $\Pi_i$  that contains, or passes close to, the mechanical axis and the optical axis of the source camera, by finding and back-projecting the line  $\mathbf{n}_i$  where  $\Pi_i$  intersects the source image  $I_i$ ; (iv) defining a 3D motion  $m$  between target and source cameras as the rotation by an angle  $\gamma = \hat{\beta} - \beta$  around a direction  $\mathbf{n}_i$  that is the normal of the vertical plane  $\Pi_i$ , such that  $\mathbf{x} = m_{\mathbf{x}}; \gamma, \mathbf{n}_i$ , with  $\hat{\mathbf{x}}$  being a point in the canonical

image of the target camera; (v) computing a focal length  $\hat{f}$  and location of the principal point  $\hat{\mathbf{O}}_i$  for the target camera and deriving the corresponding camera model that maps points  $\hat{\mathbf{x}}$  into points  $\hat{\mathbf{u}}$  in the pixel image according to  $\hat{\mathbf{u}} = c_t(\hat{\mathbf{x}}; \hat{f}, \hat{\mathbf{O}}_i, \hat{\xi})$ ; and (vi) generating a target image  $\hat{I}_i$  using image warping techniques where each pixel  $\hat{\mathbf{u}}$  in  $\hat{I}_i$  is mapped into a point  $\mathbf{u}$  in the source image  $I_i$  by a mapping function  $w$  that is the composition of the functions  $c_s$ ,  $m$  and the inverse of  $c_t$  computed in step

(v) ( $W = c_s \circ m \circ c_t^{-1}$ ), such that the color value of  $\hat{\mathbf{u}}$  can be interpolated.

**[0089]** In an embodiment of the first aspect, the azimuth  $\alpha_0$  is referenced by a particular notch position  $\mathbf{P}$ , the source image  $I_i$  is processed to detect the boundary with center  $\mathbf{C}_i$  and notch position  $\mathbf{P}_i$ , the angular displacement in azimuth is estimated as  $\delta_i = \angle \mathbf{P}_i \mathbf{Q} \mathbf{P}$ , and the line  $\mathbf{n}_i$  is defined by points  $\mathbf{O}_i$  and  $\mathbf{P}_i$

**[0090]** In an embodiment of the first aspect, the method further comprises processing the rendered target image  $\hat{I}_i$  to create a black frame such that meaningful image contents are within a circular region with center  $\hat{\mathbf{C}}$  and diameter  $\hat{d}$  and to

create a notch by placing a visual mark in point  $\hat{\mathbf{P}}_i = \hat{\mathbf{C}} + \hat{d}/2 \mathbf{v}_i$  in the circular boundary, with  $\mathbf{v}_i$  being the 2D unit direction of image line  $\mathbf{n}_i$ .

**[0091]** In an embodiment of the first aspect, the region that contains meaningful image contents can take any desired geometric shape, such as, but not limited to, a conic shape, a rectangular shape, an hexagonal shape or any other polygonal shape.

**[0092]** In an embodiment of the first aspect, point  $\mathbf{Q}$  is determined on-the fly using the detection of points  $\mathbf{P}_i$  and/or  $\mathbf{C}_i$  in successive frames, in which case it does not have to be known or determined a priori.

**[0093]** In an embodiment of the first aspect, line  $\mathbf{n}_i$  is alternatively defined by points  $\mathbf{O}_i$  and  $\mathbf{Q}$  or  $\mathbf{O}_i$  and  $\mathbf{C}_i$ .

**[0094]** In an embodiment of the first aspect, the position of the principal point is given in a normalized reference frame attached to the circular boundary, in which case the computation of its pixel location  $\mathbf{O}_i$  at every frame time instant  $i$  can be accomplished without having to explicitly know the rotation center  $\mathbf{Q}$  and angular displacement in azimuth  $\delta_i$ .

**[0095]** In an embodiment of the first aspect, the source camera is equipped with an optical encoder, or any other sensing device, that measures the rotation of the scope with respect to the camera-head and estimates the angular displacement in azimuth  $\delta_i$  of step (i).

**[0096]** In an embodiment of the first aspect, the principal point  $\mathbf{O}$  is coincident with the rotation center  $\mathbf{Q}$ , in which case the camera model of the source camera does not have to be updated at every frame time instant as in step (ii).

**[0097]** In an embodiment of the first aspect, the distortion of the target camera  $\xi$  is set to zero in order for the rendered image  $\hat{I}_i$  to be distortion free.

**[0098]** In an embodiment of the first aspect, the rendering of target image  $\hat{I}_i$  in step (vi) is performed using any common method for image warping or pixel value interpolation including, but not limited to, interpolation by nearest neighbors, bilinear interpolation or bicubic interpolation.

**[0099]** In an embodiment of the first aspect, in step (v) the principal point is made coincident with the center of the boundary ( $\hat{\mathbf{O}}_i = \hat{\mathbf{C}}$ ) and the focal length is chosen such that the FoV of the target camera is the desired  $\hat{\Theta}$ , in which case the value of  $\hat{f}$  is obtained by solving  $\Phi(\hat{f}, \hat{\xi}, \hat{\Theta}/2, \hat{d}/2) = 0$ , with  $\Phi$  being the mathematical expression that relates focal length, radial distortion, image distance, and angle between back-projection rays, all of which are interdependent parameters.

**[0100]** In an embodiment of the first aspect, in step (v) the principal point is made coincident with the center of the boundary ( $\hat{\mathbf{O}}_i = \hat{\mathbf{C}}$ ) and the focal length is chosen such that the rendered image never has a region without visual content (an empty region), the method comprising: finding a location  $\bar{\mathbf{O}}_i$  in the source image where the warping function will map the principal point  $\hat{\mathbf{O}}_i$ , which can be determined by transforming the origin  $[0, 0]^T$  by a function  $g$  that is the composition of the camera model  $c_s$  and the motion  $m$  ( $g = c_s \circ m$ ); determining a limiting viewing angle  $\bar{\Psi}_i$  that is the angle between the back-projection rays of  $\bar{\mathbf{O}}_i$  and the point in the circular boundary closest to  $\bar{\mathbf{O}}_i$ ; if  $\bar{\Psi}_i < \Theta/2$ , reducing the camera FoV to  $2\bar{\Psi}_i$  and obtaining the value of  $\hat{f}$  by solving the equation  $\Phi(\hat{f}, \hat{\xi}, \bar{\Psi}_i, \hat{d}/2) = 0$ , otherwise obtaining the value of  $\hat{f}$  by solving the equation  $\Phi(\hat{f}, \hat{\xi}, \hat{\Theta}/2, \hat{d}/2) = 0$ , with  $\Phi$  being the mathematical expression that relates focal length, radial distortion, image distance, and angle between back-projection rays, all of which are interdependent parameters.

**[0101]** In an embodiment of the first aspect, in step (vi) the focal length  $\hat{f}$  and principal point  $\hat{\mathbf{O}}_i$  are such that the rendered image has no empty region and the FoV of the target camera is the specified value  $\hat{\Theta}$ , the method comprising: finding a location  $\bar{\mathbf{O}}_i$  in the source image where the warping function will map the principal point  $\hat{\mathbf{O}}_i$ , which can be determined by transforming the origin  $[0, 0]^T$  by a function  $g$  that is the composition of the camera model  $c_s$  and the motion  $m$  ( $g = c_s \circ m$ ); determining a limiting viewing angle  $\bar{\Psi}_i$  that is the angle between the back-projection rays of  $\bar{\mathbf{O}}_i$  and the point in the circular boundary closest to  $\bar{\mathbf{O}}_i$ ; if  $\bar{\Psi}_i < \hat{\Theta}/2$ , solving the following system of equations with respect to the focal length  $\hat{f}$  and  $\lambda$ :

$$\begin{cases} \Phi(\hat{f}, \hat{\xi}, \bar{\Psi}_i, \hat{d}/2 - \lambda) = 0 \\ \Phi(\hat{f}, \hat{\xi}, \hat{\Theta} - \bar{\Psi}_i, \hat{d}/2 + \lambda) = 0 \end{cases}$$

and obtaining the principal point as  $\hat{\mathbf{O}}_i = \hat{\mathbf{C}} + \lambda \mathbf{v}_i$ , otherwise making  $\hat{\mathbf{O}}_i = \hat{\mathbf{C}}$  and finding  $\hat{f}$  by solving  $\Phi(\hat{f}, \hat{\xi}, \hat{\Theta}/2, \hat{d}/2) = 0$ , with  $\Phi$  being the mathematical expression that relates focal length, radial distortion, image distance, and angle between back-projection rays, all of which are interdependent parameters.

[0102] In an embodiment of the first aspect, the parameters of the target camera can be arbitrary, be predefined to accomplish a certain purpose, be equal to the calibration of a particular real camera equipped with a particular real endoscope with cut angle  $\hat{\beta}$ , be chosen by the user at the start of the procedure, or vary during operation according to particular events or user commands.

[0103] In an embodiment of the first aspect, the method is used in a system connected to a source camera for the purpose of empowering the user with the possibility of electronically switching between two virtual target cameras with lens cuts and FoVs of  $\hat{\beta}_1, \hat{\Theta}_1$  and  $\hat{\beta}_2, \hat{\Theta}_2$ .

[0104] In an embodiment of the first aspect, the source camera is chosen such that the cut angle of the rigid endoscope is

approximately  $\beta = \frac{\hat{\beta}_1 + \hat{\beta}_2}{2} + \frac{\hat{\Theta}_1 - \hat{\Theta}_2}{4}$  and the FoV is approximately  $\Theta = \frac{\hat{\Theta}_1 + \hat{\Theta}_2}{2} + \hat{\beta}_2 - \hat{\beta}_1$  to assure that the principal point in the target cameras is close to the center of the image.

[0105] In an embodiment of the first aspect, the FoV  $\hat{\Theta}$ , the distortion  $\hat{\xi}$ , the image resolution and the diameter  $\hat{d}$  of the target camera are set with the same values as the source camera's and the cut angle  $\hat{\beta}$  is set by the user that controls the amount of angular shift  $\gamma$  that is added to the lens cut  $\beta$  of the source camera such that  $\beta = \beta + \gamma$ .

[0106] In an embodiment of the first aspect, both the distortion  $\hat{\xi}$  and the angular shift  $\gamma$  are set to zero, in which case the system will correct the radial distortion in the source image.

[0107] In an embodiment of the first aspect, the distortion  $\hat{\xi}$  is set to zero for the target image to be rendered with no radial distortion independently of the chosen angular shift  $\gamma$ .

[0108] In an embodiment of the first aspect, the angular shift  $\gamma$  is chosen such that the principal point is placed in a region of interest in the target image and the radial distortion parameter  $\xi$  is increased to magnify this region of interest while maintaining the FoV of the target camera and all contents visible.

[0109] In an embodiment of the first aspect, the source and/or target cameras have radial distortion described by any of the models known in the literature, including, but not limited to, learning-based models, Brown's polynomial model, the rational model, the fish-eye model, or the division model.

[0110] In an embodiment of the first aspect, the boundary with center  $\mathbf{C}_i$  and the notch  $\mathbf{P}_i$  are detected using a generic conic detection method, using machine learning or deep learning techniques, or any other image processing technique.

[0111] In an embodiment of the first aspect, multiple notch positions are detected and a particular notch of these multiple notches is used for determining the angular displacement in azimuth  $\delta_i$  as the angle defined by points  $\mathbf{P}_i, \mathbf{Q}, \mathbf{P}$ .

[0112] In an embodiment of the first aspect, the Field-of-View  $\hat{\Theta}$  and the distortion  $\hat{\xi}$  of the virtual target camera take the same values as the source camera's, but the lens cut  $\hat{\beta}$  is set by the user at every frame time instant.

[0113] In an embodiment of the first aspect, the Field-of-View  $\hat{\Theta}$  of the virtual target camera takes the same value as the source camera's, but the lens cut  $\hat{\beta}$  and the distortion  $\xi$  is set by the user at every frame time instant.

[0114] In an embodiment of the first aspect, the Field-of-View  $\hat{\Theta}$  and the lens cut  $\hat{\beta}$  of the virtual target camera take the same values as the source camera's, but the distortion  $\xi$  is set by the user at every frame time instant to produce a zoom in or zoom out effect that does not change the FoV.

[0115] In an embodiment of the first aspect, the Field-of-View  $\hat{\Theta}$  of the virtual target camera takes the same value as the source camera's, but the lens cut  $\hat{\beta}$  and the distortion  $\xi$  are set by the user at every frame time instant to adapt the DoV and produce a zoom in or zoom out effect that does not change the FoV.

[0116] In an embodiment of the first aspect, the Field-of-View  $\hat{\Theta}$  of the virtual target camera takes the same value as the source camera's, but the lens cut  $\hat{\beta}$  and the focal length  $\hat{f}$  are set by the user at every frame time instant, in which case the variable to solve for is not the focal length  $\hat{f}$  but distortion  $\hat{\xi}$  to produce a zoom in or zoom out effect that does not change the FoV.

[0117] It will be appreciated that several of the above-disclosed and other features and functions, or alternatives thereof, may be desirably combined into many other different systems or applications. All such modifications and variations are intended to be included herein within the scope of this disclosure, as fall within the scope of the appended claims.

## Claims

1. A computer-implemented method for rendering a target image  $\hat{I}_i$  of a virtual target camera based on a source image  $I_i$  captured by a real source camera, the source camera comprising a camera-head and a rigid endoscope with lens cut  $\beta$ , the lens cut being the angle between the optical and mechanical axes of the endoscope, that rotates in azimuth around a mechanical axis that intersects the image in a point  $\mathbf{Q}$ , for which the focal length  $f$ , radial distortion  $\xi$ , and principal point  $\mathbf{O}$  at a certain azimuth  $\alpha_0$  are known, the method comprising:

finding a location of the principal point  $\mathbf{O}_i$  in the source image  $I_i$  with azimuth  $\alpha_i$  by rotating  $\mathbf{O}$  around  $\mathbf{Q}$  by an angular displacement in azimuth  $\delta_i = \alpha_i - \alpha_0$ ;

updating a camera model  $c_s$  of the source camera that maps points  $x$  in a canonical image into points  $u$  in a pixel image according to the focal length  $f$ , radial distortion  $\xi$ , and location of the principal point  $O_i$ ;  
determining a 3D location of a vertical plane  $\Pi_i$  that contains the mechanical axis and the optical axis of the source camera, by finding and back-projecting into 3D space a line  $n_i$  where  $\Pi_i$  intersects the source image  $I_i$ ; defining a  
5 3D motion  $m$  between the target and source cameras as a rotation by an angle  $\gamma = \hat{\beta} - \beta$  around a direction  $\vec{n}_i$  that is normal to the vertical plane  $\Pi_i$ , such that  $x = m(\hat{x}; \gamma, \vec{n}_i)$ , with  $\hat{x}$  being a point in the canonical image of the target camera and  $\hat{\beta}$  being the lens cut of the target camera;  
computing a focal length  $\hat{f}$  and a location of the principal point  $\hat{O}_i$  for the target camera and deriving a camera  
10 model  $c_t$  that maps the points  $\hat{x}$  in the canonical image of the target image into points  $\hat{u}$  in a pixel image of the target image according to the focal length  $\hat{f}$  and location of the principal point  $\hat{O}_i$ ; and  
generating the target image  $\hat{I}_i$  by mapping a plurality of pixels  $\hat{u}$  in  $\hat{I}_i$  into a point  $u$  in the source image  $I_i$  with a mapping function  $w$  that is a composition of the camera model  $c_s$  of the source camera, the 3D motion  $m$ , and an inverse of a camera model  $c_t$  of the target camera.

15 2. The method of claim 1, wherein the azimuth  $\alpha_0$  corresponds to a first notch position  $P$ , the method further comprising:

processing the source image  $I_i$  to detect a boundary with a center  $C_i$  and a second notch position  $P_i$ ;  
wherein rotating  $O$  around  $Q$  by an angular displacement in azimuth comprises estimating the angular displacement in azimuth according to the first notch position  $P$ , the second notch position  $P_i$ , and the point  $Q$ ;  
20 wherein the line  $n_i$  is defined by points  $O_i$  and  $P_i$ .

3. The method in claim 2, further comprising processing the rendered target image  $\hat{I}_i$  to create a black frame defining an

image region with center  $\hat{C}$  and diameter  $\hat{d}$  and to create a notch by placing a visual mark at point  $\hat{P}_i = \hat{C} + \hat{d}/2 v_i$   
25 in the circular boundary, with  $v_i$  being the 2D unit direction of image line  $n_i$ .

4. The method of claim 3, wherein the image region comprises one or more of a circular shape, a conic shape, a rectangular shape, an hexagonal shape, or another polygonal shape.

30 5. The method of claim 2, wherein the source image  $I_i$  comprises two or more source image frames, wherein point  $Q$  is determined by detecting one or more of point  $P_i$  or point  $C_i$  in successive ones of the two or more frames.

6. The method of claim 2, wherein processing the source image  $I_i$  to detect a boundary with a center  $C_i$  and a second notch position  $P_i$  comprises detecting a plurality of second notch positions  $P_i$ , wherein one of the plurality of second notch positions is used for determining the angular displacement in azimuth  $\delta_i$ .  
35

7. The method of claim 1, wherein the azimuth  $\alpha_0$  corresponds to a first notch position  $P$ , the method further comprising:

processing the source image  $I_i$  to detect a boundary with a center  $C_i$  and a second notch position  $P_i$ ;  
wherein rotating  $O$  around  $Q$  by an angular displacement in azimuth comprises estimating the angular displacement in azimuth according to the first notch position  $P$ , the second notch position  $P_i$ , and the point  $Q$ ;  
40 wherein line  $n_i$  is defined by:

45 points  $O_i$  and  $Q$ ; or  
points  $O_i$  and  $C_i$ .

8. The method of claim 1, wherein the position of the principal point is given in a normalized reference frame disposed in or attached to the endoscope.

50 9. The method of claim 1, wherein the source camera is equipped with a sensor that measures the rotation of the endoscope with respect to the camera-head and estimates the angular displacement in azimuth  $\delta_i$ .

10. The method of claim 1, wherein the principal point  $O$  is coincident with the rotation center  $Q$ .

55 11. The method of claim 1, wherein a distortion of the target camera  $\xi$  is set to zero.

12. The method of claim 1, wherein generating the target image  $\hat{I}_i$  is performed using an image warping or pixel value interpolation comprising one or more of interpolation by nearest neighbors, bilinear interpolation, or bicubic inter-

polation.

13. The method of claim 1, wherein the principal point  $\hat{O}_i$  is made coincident with a center  $\hat{C}$  of a boundary of the virtual image and computing the focal length  $f$  comprises solving  $\Phi(\hat{f}, \hat{\xi}, \hat{\theta}/2, \hat{d}/2) = 0$ , with  $\Phi$  being a mathematical expression that relates focal length, radial distortion, image distance, and angle between back-projection rays.

14. The method of claim 1 wherein the principal point  $\hat{O}_i$  is made coincident with a center  $\hat{C}$  of a boundary of the virtual image, the method comprising:

transforming an origin  $[0, 0]^T$  of the source image by a function  $g$  that is the composition of the camera model  $c_s$  and the motion  $m$  to find a location  $\bar{O}_i$  in the source image that maps to the principal point  $\hat{O}_i$ ;  
determining a limiting viewing angle  $\bar{\Psi}_i$  between the back-projection rays of  $\bar{O}_i$  and a point in the circular boundary closest to  $\bar{O}_i$ ;

when  $\bar{\Psi}_i < \hat{\theta}/2$ , reducing the camera field of view to  $2\bar{\Psi}_i$  and obtaining the value of  $\hat{f}$  by solving the equation  $\Phi(\hat{f}, \hat{\xi}, \bar{\Psi}_i, \hat{d}/2) = 0$ , with  $\Phi$  being a mathematical expression that relates focal length, radial distortion, image distance, and angle between back-projection rays;

when  $\bar{\Psi}_i \geq \hat{\theta}/2$ , obtaining the value of  $f$  by solving the equation  $\Phi(\hat{f}, \hat{\xi}, \hat{\theta}/2, \hat{d}/2) = 0$ , with  $\Phi$  being a mathematical expression that relates focal length, radial distortion, image distance, and angle between back-projection rays, all of which are interdependent parameters.

15. The method of claim 1, further comprising:

transforming an origin  $[0, 0]^T$  of the source image by a function  $g$  that is the composition of the camera model  $c_s$  and the motion  $m$  to find a location  $\bar{O}_i$  in the source image that maps to the principal point  $\hat{O}_i$ ;  
determining a limiting viewing angle  $\bar{\Psi}_i$  between the back-projection rays of  $\bar{O}_i$  and a point in the circular boundary closest to  $\bar{O}_i$ ;

when  $\bar{\Psi}_i < \hat{\theta}/2$ , solving the following system of equations with respect to the focal length  $\hat{f}$  and  $\lambda$ :

$$\begin{cases} \Phi(\hat{f}, \hat{\xi}, \bar{\Psi}_i, \hat{d}/2 - \lambda) = 0 \\ \Phi(\hat{f}, \hat{\xi}, \hat{\theta} - \bar{\Psi}_i, \hat{d}/2 + \lambda) = 0 \end{cases}$$

, and obtaining the principal point as  $\hat{O}_i = \hat{C} + \lambda \mathbf{v}_i$ ;

when  $\bar{\Psi}_i \geq \hat{\theta}/2$ , setting  $\hat{O}_i = \hat{C}$  and determining  $f$  by solving  $\Phi(\hat{f}, \hat{\xi}, \hat{\theta}/2, \hat{d}/2) = 0$ , with  $\Phi$  being the mathematical expression that relates focal length, radial distortion, image distance, and angle between back-projection rays, all of which are interdependent parameters.

## Patentansprüche

1. Computerimplementiertes Verfahren zum Rendern eines Zielbildes  $\hat{I}_i$  einer virtuellen Zielkamera auf der Basis eines von einer realen Quellkamera aufgenommenen Quellbildes  $I_i$ , wobei die Quellkamera einen Kamerakopf und ein starres Endoskop mit einem Linsenschnitt  $\beta$  umfasst, wobei der Linsenschnitt der Winkel zwischen der optischen und der mechanischen Achse des Endoskops ist, das sich im Azimut um eine mechanische Achse dreht, die das Bild in einem Punkt  $Q$  schneidet, für den die Brennweite  $f$ , die radiale Verzerrung  $\xi$  und der Hauptpunkt  $O$  bei einem bestimmten Azimut  $\alpha_0$  bekannt sind, wobei das Verfahren Folgendes umfasst:

Finden einer Position des Hauptpunkts  $O_i$  im Quellbild  $I_i$  mit Azimut  $\alpha_i$  durch Drehen von  $O$  um  $Q$  um eine Winkelverschiebung im Azimut  $\delta_i = \alpha_i - \alpha_0$ ;

Aktualisieren eines Kameramodells  $c_s$  der Quellkamera, das Punkte  $x$  in einem kanonischen Bild in Punkte  $u$  in einem Pixelbild gemäß der Brennweite  $f$ , der radialen Verzerrung  $\xi$  und der Position des Hauptpunkts  $O_i$  abbildet;  
Bestimmen einer 3D-Position einer vertikalen Ebene  $\Pi_i$ , die die mechanische Achse und die optische Achse der Quellkamera enthält, durch Auffinden und Rückprojizieren einer Linie  $n_i$  in den 3D-Raum, wobei  $\Pi_i$  das Quellbild  $I_i$  schneidet;

Definieren einer 3D-Bewegung  $m$  zwischen Ziel- und Quellkamera als eine Drehung um einen Winkel  $\gamma = \hat{\beta} - \beta$  um eine Richtung  $\vec{n}_i$ , die normal zur vertikalen Ebene  $\Pi_i$  ist, so dass  $x = m(\hat{x}; \gamma, \vec{n}_i)$ , wobei  $\hat{x}$  ein Punkt im kanonischen Bild der Zielkamera und der Linsenschnitt der Zielkamera ist;

Berechnen einer Brennweite  $\hat{f}$  und einer Position des Hauptpunkts  $\hat{O}_i$  für die Zielkamera und Ableiten eines Kameramodells  $c_t$ , das die Punkte  $\hat{x}$  im kanonischen Bild des Zielbildes gemäß der Brennweite  $\hat{f}$  und der Position des Hauptpunktes  $\hat{O}_i$  in Punkte  $\hat{u}$  in einem Pixelbild des Zielbildes abbildet; und

Erzeugen des Zielbildes  $\hat{I}_i$  durch Abbilden mehrerer Pixel  $\hat{u}$  in  $\hat{I}_i$  auf einen Punkt  $u$  im Quellbild  $I_i$  mit einer Abbildungsfunktion  $w$ , die eine Zusammensetzung des Kameramodells  $c_s$  der Quellkamera, der 3D-Bewegung  $m$  und einer Inversen eines Kameramodells  $c_t$  der Zielkamera ist.

- 5 2. Verfahren nach Anspruch 1, wobei der Azimut  $\alpha_0$  einer ersten Kerbenposition  $P$  entspricht, wobei das Verfahren ferner Folgendes umfasst:

Verarbeiten des Quellbildes  $I_i$ , um eine Grenze mit einem Mittelpunkt  $C_i$  und einer zweiten Kerbenposition  $P_i$  zu erfassen;

- 10 wobei das Drehen von  $O$  um  $Q$  um eine Winkelverschiebung im Azimut das Schätzen der Winkelverschiebung im Azimut gemäß der ersten Kerbenposition  $P$ , der zweiten Kerbenposition  $P_i$  und dem Punkt  $Q$  umfasst; wobei die Linie  $n_i$  durch die Punkte  $O_i$  und  $P_i$  definiert ist.

- 15 3. Verfahren nach Anspruch 2, das ferner das Verarbeiten des gerenderten Zielbildes  $\hat{I}_i$  umfasst, um einen schwarzen Rahmen zu erzeugen, der eine Bildregion mit Mittelpunkt  $\hat{C}$  und Durchmesser  $\hat{d}$  definiert, und um eine Kerbe zu

erzeugen, indem eine visuelle Markierung an Punkt  $\hat{P}_i = \hat{C} + \hat{d}/2 v_i$  in der kreisförmigen Begrenzung platziert wird, wobei  $v_i$  die 2D-Einheitsrichtung der Bildlinie  $n_i$  ist.

- 20 4. Verfahren nach Anspruch 3, wobei die Bildregion eines von einer Kreisform, einer Kegelform, einer Rechteckform, einer Sechseckform und einer anderen polygonalen Form hat.

- 25 5. Verfahren nach Anspruch 2, wobei das Quellbild  $I_i$  zwei oder mehr Quellbildrahmen umfasst, wobei der Punkt  $Q$  durch Erkennen von Punkt  $P_i$  und/oder  $C_i$  in aufeinanderfolgenden der zwei oder mehr Rahmen bestimmt wird.

- 30 6. Verfahren nach Anspruch 2, wobei die Verarbeitung des Quellbildes  $I_i$  zum Erkennen einer Grenze mit einem Mittelpunkt  $C_i$  und einer zweiten Kerbenposition  $P_i$  das Erkennen mehrerer zweiter Kerbenpositionen  $P_i$  umfasst, wobei eine der mehreren zweiten Kerbenpositionen zum Bestimmen der Winkelverschiebung im Azimut  $\delta_i$  verwendet wird.

7. Verfahren nach Anspruch 1, wobei der Azimut  $\alpha_0$  einer ersten Kerbenposition  $P$  entspricht, wobei das Verfahren ferner Folgendes umfasst:

35 Verarbeiten des Quellbildes  $I_i$ , um eine Grenze mit einem Mittelpunkt  $C_i$  und einer zweiten Kerbenposition  $P_i$  zu erfassen;

wobei das Drehen von  $O$  um  $Q$  um eine Winkelverschiebung im Azimut das Schätzen der Winkelverschiebung im Azimut gemäß der ersten Kerbenposition  $P$ , der zweiten Kerbenposition  $P_i$  und dem Punkt  $Q$  umfasst; wobei die Linie  $n_i$  definiert ist durch:

40 Punkte  $O_i$  und  $Q_i$ ; oder  
Punkte  $O_i$  und  $C_i$ .

- 45 8. Verfahren nach Anspruch 1, wobei die Position des Hauptpunkts in einem normalisierten Referenzrahmen angegeben ist, der in dem Endoskop angeordnet oder daran angebracht ist.

9. Verfahren nach Anspruch 1, wobei die Quellkamera mit einem Sensor ausgestattet ist, der die Drehung des Endoskops in Bezug auf den Kamerakopf misst und die Winkelverschiebung im Azimut  $\delta_i$  schätzt.

- 50 10. Verfahren nach Anspruch 1, wobei der Hauptpunkt  $O$  mit dem Drehmittelpunkt  $Q$  zusammenfällt.

11. Verfahren nach Anspruch 1, wobei eine Verzerrung der Zielkamera  $\xi$  auf Null gesetzt wird.

- 55 12. Verfahren nach Anspruch 1, wobei das Erzeugen des Zielbildes  $\hat{I}_i$  mittels einer Bildverzerrung oder Pixelwertinterpolation durchgeführt wird, die eines oder mehrere von Interpolation durch nächste Nachbarn, bilinearer Interpolation und bikubischer Interpolation umfasst.

13. Verfahren nach Anspruch 1, wobei der Hauptpunkt  $\hat{O}_i$  mit einem Mittelpunkt  $C$  einer Grenze des virtuellen Bildes in Übereinstimmung gebracht wird und das Berechnen der Brennweite  $\hat{f}$  das Lösen von  $\Phi(\hat{f}, \hat{\xi}, \hat{\theta}/2, \hat{d}/2) = 0$  umfasst,

wobei  $\Phi$  ein mathematischer Ausdruck ist, der Brennweite, radiale Verzerrung, Bildabstand und Winkel zwischen Rückprojektionsstrahlen in Beziehung setzt.

- 5 14. Verfahren nach Anspruch 1, wobei der Hauptpunkt  $\hat{O}_i$  mit einem Mittelpunkt  $\hat{C}$  einer Grenze des virtuellen Bildes in Übereinstimmung gebracht wird, wobei das Verfahren Folgendes umfasst:

Transformieren eines Ursprungs  $[0, 0]^T$  des Quellbildes durch eine Funktion  $g$ , die die Zusammensetzung des Kameramodells  $c_s$  und der Bewegung  $m$  ist, um eine Position  $\bar{O}_i$  im Quellbild zu finden, die auf den Hauptpunkt  $\hat{O}_i$  abgebildet ist;

10 Bestimmen eines begrenzenden Betrachtungswinkels  $\bar{\Psi}_i$  zwischen den Rückprojektionsstrahlen von  $\bar{O}_i$  und einem Punkt in der kreisförmigen Begrenzung, der  $\bar{O}_i$  am nächsten liegt;

Reduzieren, wenn  $\bar{\Psi}_i < \hat{\Theta}/2$ , des Kamerasichtfeldes auf  $2\bar{\Psi}_i$  und Erhalten des Wertes von  $\hat{f}$  durch Lösen der Gleichung  $\Phi(\hat{f}, \hat{\xi}, \bar{\Psi}_i, \hat{d}/2) = 0$ , wobei  $\Phi$  ein mathematischer Ausdruck ist, der Brennweite, radiale Verzerrung, Bildabstand und Winkel zwischen Rückprojektionsstrahlen in Beziehung setzt;

15 Erhalten, wenn  $\bar{\Psi}_i > \hat{\Theta}/2$ , des Wertes von  $\hat{f}$  durch Lösen der Gleichung  $\Phi(\hat{f}, \hat{\xi}, \hat{\Theta}/2, \hat{d}/2) = 0$ , wobei  $\Phi$  ein mathematischer Ausdruck ist, der Brennweite, radiale Verzerrung, Bildabstand und Winkel zwischen Rückprojektionsstrahlen in Beziehung setzt, wobei alle diese Parameter voneinander abhängen.

- 20 15. Verfahren nach Anspruch 1, das ferner Folgendes umfasst:

Transformieren eines Ursprungs  $[0, 0]^T$  des Quellbildes durch eine Funktion  $g$ , die die Zusammensetzung des Kameramodells  $c_s$  und der Bewegung  $m$  ist, um eine Position  $\bar{O}_i$  im Quellbild zu finden, die auf den Hauptpunkt  $\hat{O}_i$  abgebildet ist;

25 Bestimmen eines begrenzenden Betrachtungswinkels  $\bar{\Psi}_i$  zwischen den Rückprojektionsstrahlen von  $\bar{O}_i$  und einem Punkt in der kreisförmigen Begrenzung, der  $\bar{O}_i$  am nächsten liegt;

Lösen, wenn  $\bar{\Psi}_i < \hat{\Theta}/2$ , des folgenden Gleichungssystems in Bezug auf die Brennweite  $\hat{f}$  und  $\lambda$ :

$$\begin{cases} \Phi(\hat{f}, \hat{\xi}, \bar{\Psi}_i, \hat{d}/2 - \lambda) = 0 \\ \Phi(\hat{f}, \hat{\xi}, \hat{\Theta} - \bar{\Psi}_i, \hat{d}/2 + \lambda) = 0 \end{cases}$$

und Erhalten des Hauptpunkts als  $\hat{O}_i = \hat{C} + \lambda \mathbf{v}_i$ ,

30 wenn  $\bar{\Psi}_i > \hat{\Theta}/2$ , Setzen von  $\hat{O}_i = C$  und Bestimmen von  $\hat{f}$  durch Lösen von  $\Phi(\hat{f}, \hat{\xi}, \hat{\Theta}/2, \hat{d}/2) = 0$ , wobei  $\Phi$  der mathematische Ausdruck ist, der Brennweite, radiale Verzerrung, Bildabstand und Winkel zwischen Rückprojektionsstrahlen in Beziehung setzt, die alle voneinander abhängige Parameter sind.

## 35 Revendications

- 40 1. Procédé informatique pour le rendu d'une image cible  $\hat{I}_i$  d'une caméra cible virtuelle sur la base d'une image source li capturée par une caméra source réelle, la caméra source comprenant une tête de caméra et un endoscope rigide avec une déviation de lentille  $\beta$ , la déviation de lentille étant l'angle entre les axes optique et mécanique de l'endoscope, qui tourne en azimut autour d'un axe mécanique qui coupe l'image en un point Q, pour lequel la distance focale  $f$ , la distorsion radiale  $\xi$  et le point principal O à un certain azimut  $\alpha_0$  sont connus, le procédé consistant à :

45 trouver un emplacement du point principal  $O_i$  sur l'image source li avec l'azimut  $\alpha_i$  en faisant tourner O autour de Q d'un déplacement angulaire en azimut  $\delta_i = \alpha_i - \alpha_0$  ;

mettre à jour un modèle de caméra  $c_s$  de la caméra source qui met en correspondance des points  $x$  sur une image canonique avec des points  $u$  sur une image pixellique selon la distance focale  $f$ , la distorsion radiale  $\xi$  et l'emplacement du point principal  $O_i$  ;

50 déterminer un emplacement 3D d'un plan vertical  $\Pi_i$  qui contient l'axe mécanique et l'axe optique de la caméra source, en trouvant et en rétroprojetant dans l'espace 3D une ligne  $n_i$  où  $\Pi_i$  coupe l'image source li ;

définir un mouvement 3D  $m$  entre les caméras cible et source comme une rotation d'un angle  $\gamma = \hat{\beta} - \beta$  autour d'une direction  $\vec{n}_i$  qui est normale au plan vertical  $\Pi_i$ , de telle sorte que  $x = m(\hat{x}; \gamma, \vec{n}_i)$ ,  $\hat{x}$  étant un point sur l'image canonique de la caméra cible et étant la déviation de lentille de la caméra cible ;

55 calculer une distance focale  $\hat{f}$  et un emplacement du point principal  $\hat{O}_i$  pour la caméra cible et dériver un modèle de caméra  $c_t$  qui met en correspondance les points  $\hat{x}$  sur l'image canonique de l'image cible avec des points  $\hat{u}$  sur une image pixellique de l'image cible selon la distance focale  $\hat{f}$  et l'emplacement du point principal  $\hat{O}_i$  ; et

générer l'image cible  $\hat{I}_i$  en mettant en correspondance une pluralité de pixels  $\hat{u}$  sur  $\hat{I}_i$  avec un point  $u$  sur l'image

source  $I_i$  à l'aide d'une fonction de correspondance  $w$  qui est une composition du modèle de caméra  $c_s$  de la caméra source, du mouvement 3D  $m$  et d'un inverse d'un modèle de caméra  $c_t$  de la caméra cible.

2. Procédé selon la revendication 1, dans lequel l'azimut  $\alpha_0$  correspond à une première position d'encoche P, le procédé consistant en outre à :

traiter l'image source  $I_i$  pour détecter une limite avec un centre  $C_i$  et une seconde position d'encoche  $P_i$  ;  
le fait de faire tourner O autour de Q d'un déplacement angulaire en azimut consistant à estimer le déplacement angulaire en azimut selon la première position d'encoche P, la seconde position d'encoche  $P_i$  et le point Q ;  
la ligne  $n_i$  étant définie par les points  $O_i$  et  $P_i$ .

3. Procédé selon la revendication 2, consistant en outre à traiter l'image cible  $\hat{I}_i$  rendue pour créer un cadre noir définissant une région d'image avec un centre C et un diamètre  $\hat{d}$  et pour créer une encoche en plaçant une marque

visuelle à un point  $\hat{P}_i = \hat{C} + \hat{d}/2 v_i$  sur la limite circulaire,  $v_i$  étant la direction unitaire 2D de la ligne d'image  $n_i$ .

4. Procédé selon la revendication 3, dans lequel la région d'image comprend une ou plusieurs formes parmi une forme circulaire, une forme conique, une forme rectangulaire, une forme hexagonale et une autre forme polygonale.

5. Procédé selon la revendication 2, dans lequel l'image source  $I_i$  comprend deux trames d'image source ou plus, le point Q étant déterminé en détectant le point  $P_i$  ou le point  $C_i$  dans des trames successives parmi les deux trames ou plus.

6. Procédé selon la revendication 2, dans lequel le traitement de l'image source  $I_i$  pour détecter une limite avec un centre  $C_i$  et une seconde position d'encoche  $P_i$  consiste à détecter une pluralité de secondes positions d'encoche  $P_i$ , une seconde position d'encoche parmi la pluralité de secondes positions d'encoche étant utilisée pour déterminer le déplacement angulaire en azimut  $\delta_i$ .

7. Procédé selon la revendication 1, dans lequel l'azimut  $\alpha_0$  correspond à une première position d'encoche P, le procédé consistant en outre à :

traiter l'image source  $I_i$  pour détecter une limite avec un centre  $C_i$  et une seconde position d'encoche  $P_i$  ;  
le fait de faire tourner O autour de Q d'un déplacement angulaire en azimut consistant à estimer le déplacement angulaire en azimut selon la première position d'encoche P, la seconde position d'encoche  $P_i$  et le point Q ;  
la ligne  $n_i$  étant définie par :

les points  $O_i$  et Q ; ou  
les points  $O_i$  et  $C_i$ .

8. Procédé selon la revendication 1, dans lequel la position du point principal est donnée dans un cadre de référence normalisé disposé dans l'endoscope ou attaché à celui-ci.

9. Procédé selon la revendication 1, dans lequel la caméra source est équipée d'un capteur qui mesure la rotation de l'endoscope par rapport à la tête de caméra et estime le déplacement angulaire en azimut  $\delta_i$ .

10. Procédé selon la revendication 1, dans lequel le point principal O coïncide avec le centre de rotation Q.

11. Procédé selon la revendication 1, dans lequel une distorsion de la caméra cible  $\hat{\xi}$  est fixée à zéro.

12. Procédé selon la revendication 1, dans lequel la génération de l'image cible  $\hat{I}_i$  est réalisée au moyen d'une interpolation de déformation d'image ou de valeur de pixel comprenant une ou plusieurs interpolations parmi une interpolation par plus proche voisin, une interpolation bilinéaire et une interpolation bicubique.

13. Procédé selon la revendication 1, dans lequel le point principal  $\hat{O}_i$  coïncide avec un centre  $\hat{C}$  d'une limite de l'image virtuelle et le calcul de la distance focale  $\hat{f}$  consiste à résoudre  $\Phi(\hat{f}, \hat{\xi}, \hat{\theta}/2, \hat{d}/2) = 0$ , où  $\Phi$  est une expression mathématique qui relie la distance focale, la distorsion radiale, la distance de l'image et l'angle entre les rayons de rétroprojection.

14. Procédé selon la revendication 1, dans lequel le point principal  $\hat{\mathbf{O}}_i$  est rendu coïncident avec un centre C d'une limite de l'image virtuelle, le procédé consistant à :

transformer une origine  $[0, 0]^T$  de l'image source par une fonction g qui est la composition du modèle de caméra  $c_s$  et du mouvement  $m$  pour trouver sur l'image source un emplacement  $\bar{\mathbf{O}}_i$  qui correspond au point principal  $\hat{\mathbf{O}}_i$  ;  
déterminer un angle de vision limite  $\bar{\Psi}_i$  entre les rayons de rétroprojection de  $\bar{\mathbf{O}}_i$  et un point sur la limite circulaire le plus proche de  $\bar{\mathbf{O}}_i$  ;

lorsque  $\bar{\Psi}_i < \hat{\Theta}/2$ , réduire le champ de vision de la caméra à  $2\bar{\Psi}_i$  et obtenir la valeur de  $\hat{f}$  en résolvant l'équation  $\Phi(\hat{f}, \hat{\xi}, \bar{\Psi}_i, \hat{d}/2) = 0$ , où  $\Phi$  est une expression mathématique qui relie la distance focale, la distorsion radiale, la distance de l'image et l'angle entre les rayons de rétroprojection ;

lorsque  $\bar{\Psi}_i \geq \hat{\Theta}/2$ , obtenir la valeur de  $\hat{f}$  en résolvant l'équation  $\Phi(\hat{f}, \hat{\xi}, \hat{\Theta}/2, \hat{d}/2) = 0$ , où  $\Phi$  est une expression mathématique qui relie la distance focale, la distorsion radiale, la distance de l'image et l'angle entre les rayons de rétroprojection, qui sont tous des paramètres interdépendants.

15. Procédé selon la revendication 1, consistant en outre à :

transformer une origine  $[0, 0]^T$  de l'image source par une fonction g qui est la composition du modèle de caméra  $c_s$  et du mouvement  $m$  pour trouver sur l'image source un emplacement  $\bar{\mathbf{O}}_i$  qui correspond au point principal  $\hat{\mathbf{O}}_i$  ;  
déterminer un angle de vision limite  $\bar{\Psi}_i$  entre les rayons de rétroprojection de  $\bar{\mathbf{O}}_i$  et un point sur la limite circulaire le plus proche de  $\bar{\mathbf{O}}_i$  ;

lorsque  $\bar{\Psi}_i < \hat{\Theta}/2$ , résoudre le système d'équations suivant en fonction de la distance focale  $\hat{f}$  et de  $\lambda$ :

$$\begin{cases} \Phi(\hat{f}, \hat{\xi}, \bar{\Psi}_i, \hat{d}/2 - \lambda) = 0 \\ \Phi(\hat{f}, \hat{\xi}, \hat{\Theta} - \bar{\Psi}_i, \hat{d}/2 + \lambda) = 0 \end{cases}$$

, et obtenir le point principal comme  $\hat{\mathbf{O}}_i = \hat{\mathbf{C}} + \lambda \mathbf{v}_i$  ;

lorsque  $\bar{\Psi}_i \geq \hat{\Theta}/2$ , fixer  $\hat{\mathbf{O}}_i = \hat{\mathbf{C}}$  et déterminer  $\hat{f}$  en résolvant l'équation  $\Phi(\hat{f}, \hat{\xi}, \hat{\Theta}/2, \hat{d}/2) = 0$ , où  $\Phi$  est l'expression mathématique qui relie la distance focale, la distorsion radiale, la distance de l'image et l'angle entre les rayons de rétroprojection, qui sont tous des paramètres interdépendants.

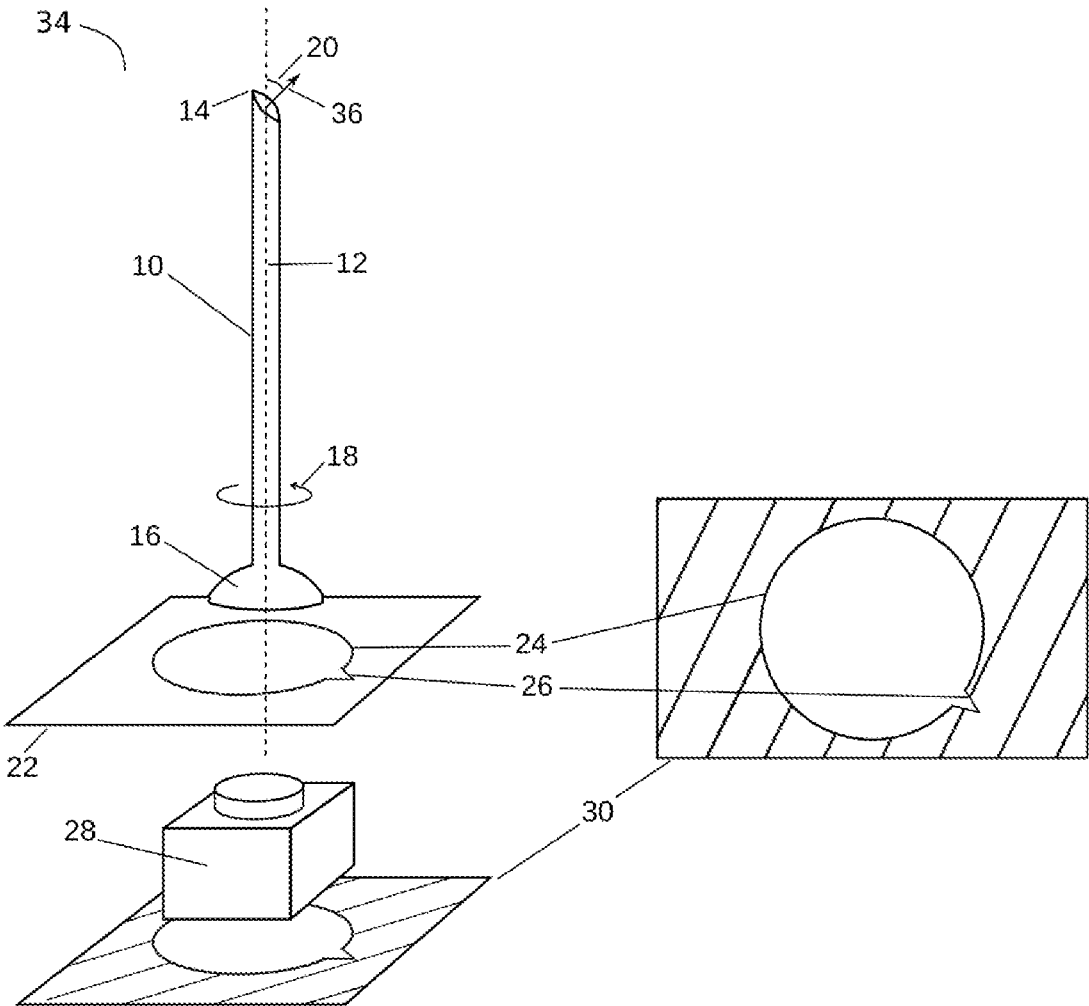


FIG. 1

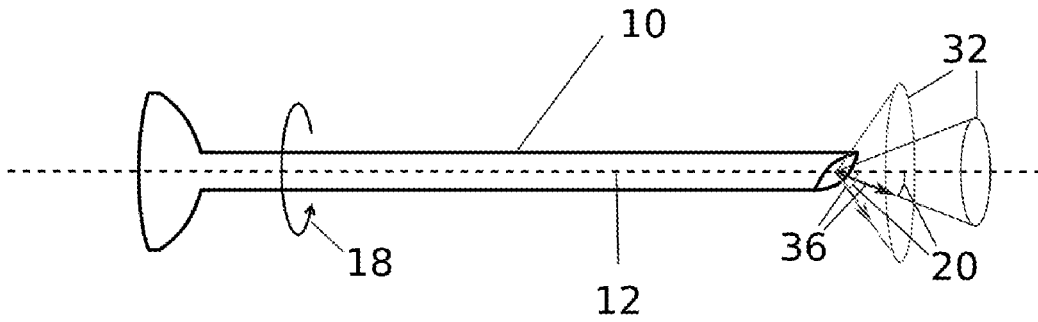
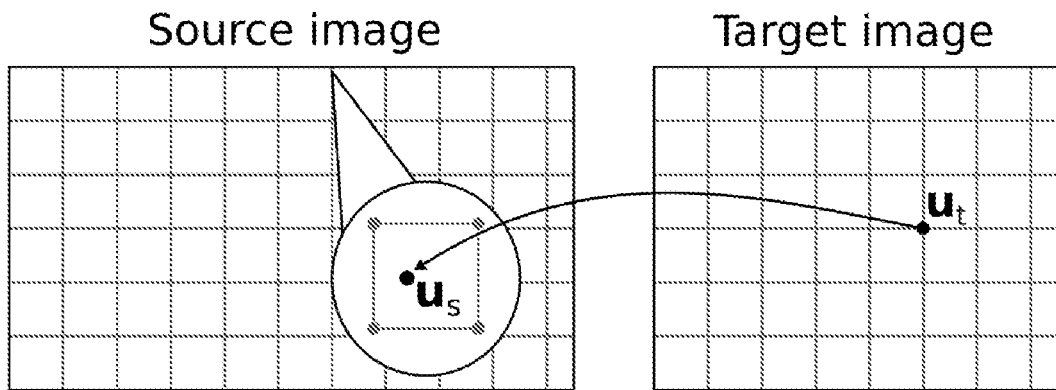


FIG. 2



$$\mathbf{u}_s = w(\mathbf{u}_t) = c_s \circ m \circ c_t^{-1}(\mathbf{u}_t)$$

FIG. 3

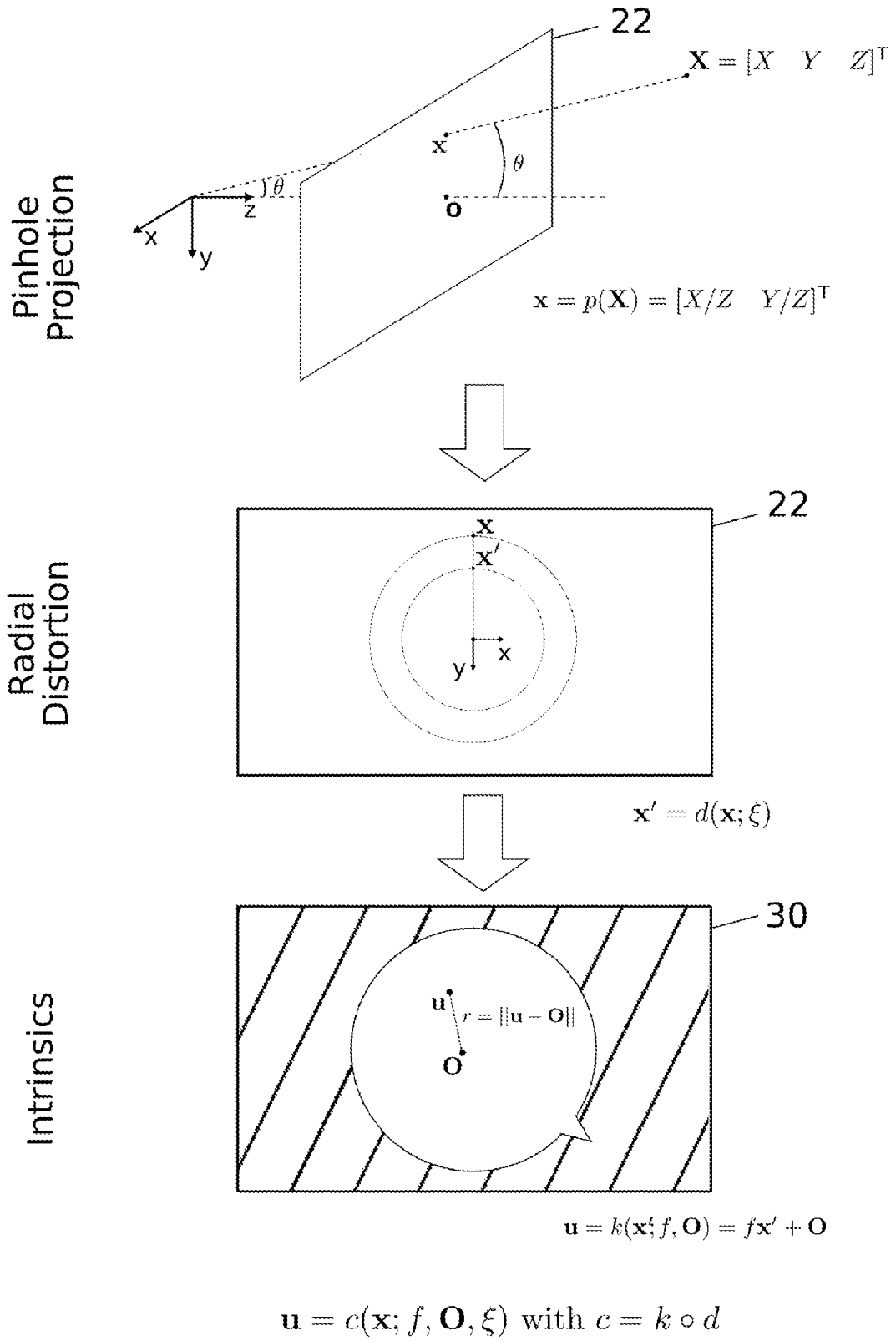


FIG. 4

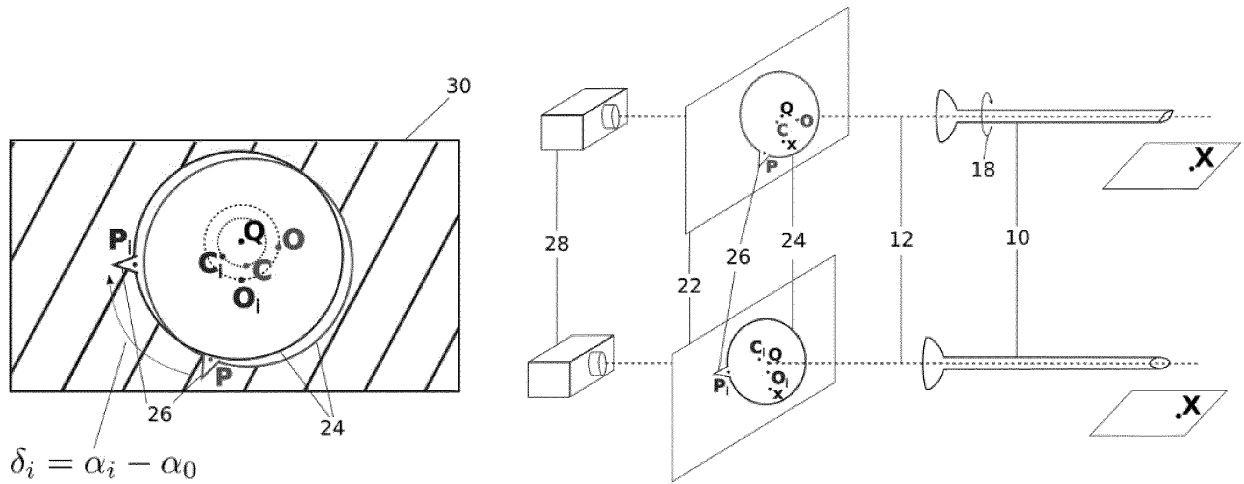


FIG. 5

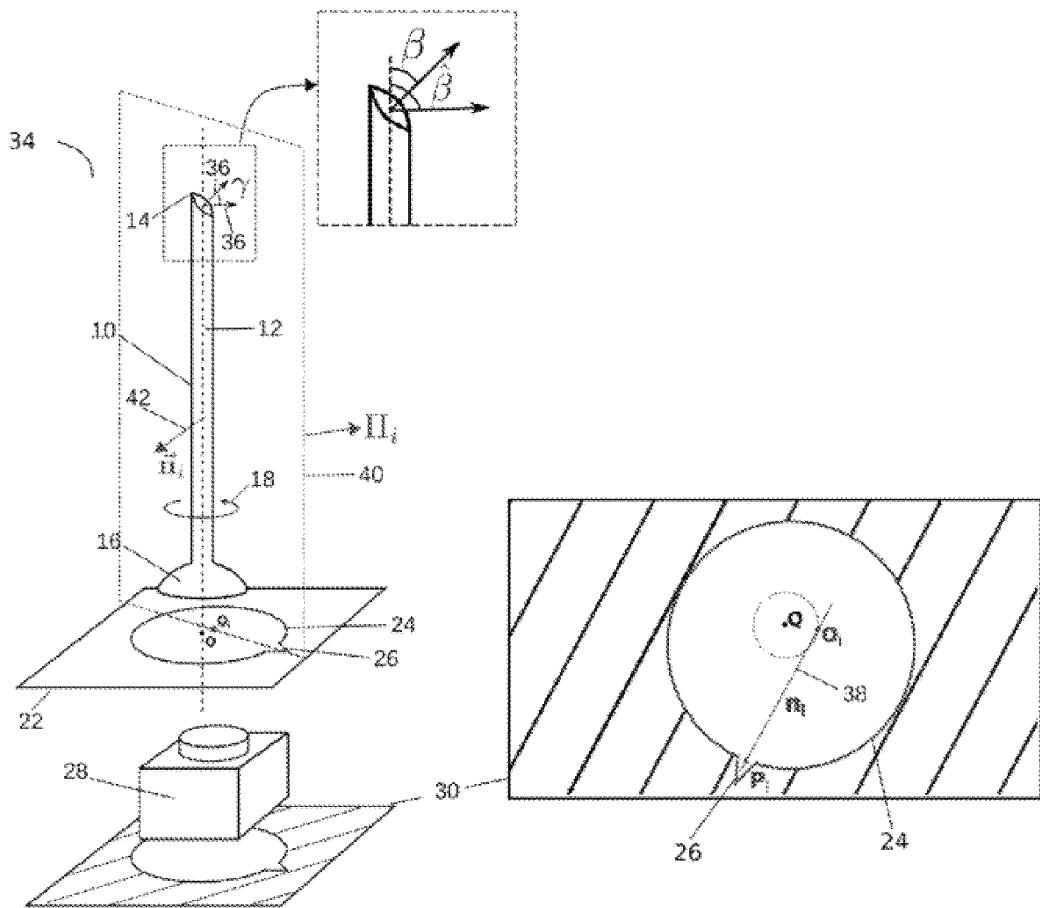


FIG. 6

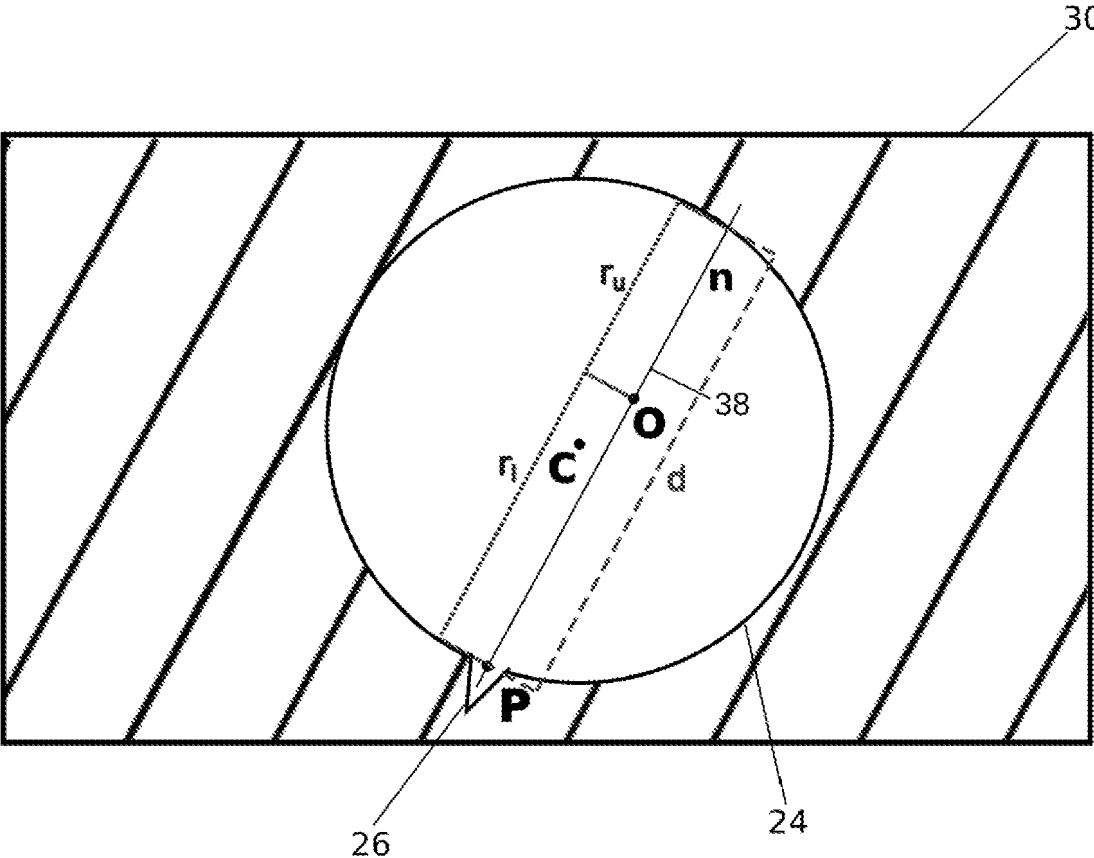


FIG. 7

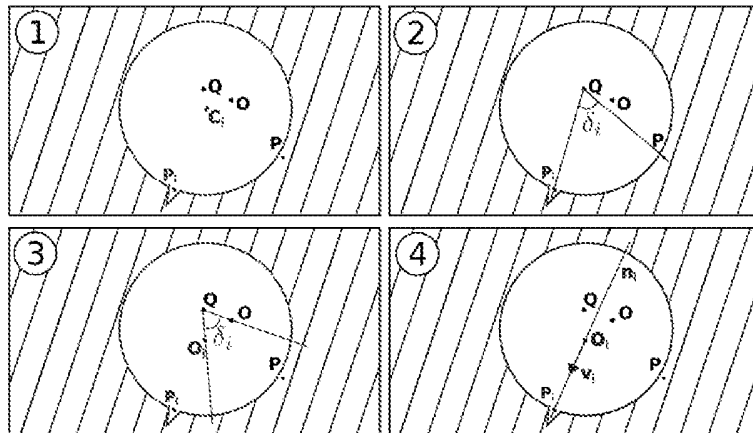
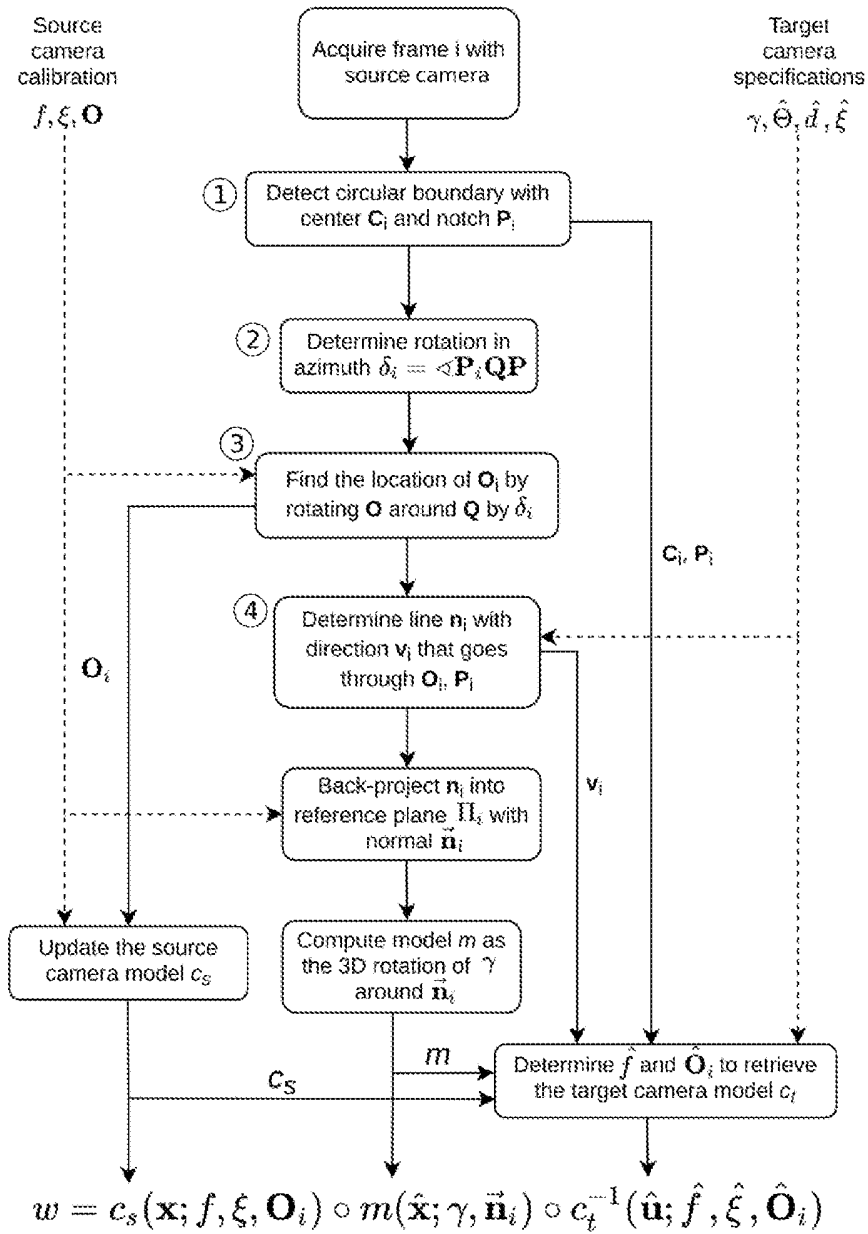
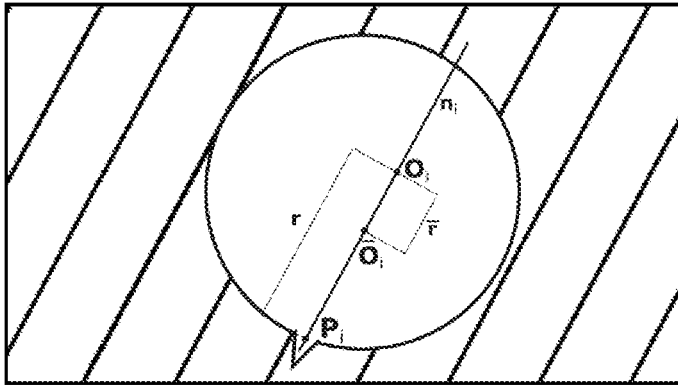


FIG. 8



Source image

$$\begin{aligned}
 \mathbf{u} &= g(\hat{\mathbf{x}}; f, \xi, \gamma, \mathbf{O}_i, \vec{n}_i) \\
 &= c_s(\mathbf{x}; f, \xi, \mathbf{O}_i) \circ m(\hat{\mathbf{x}}; \gamma, \vec{n}_i) \\
 \bar{\mathbf{O}}_i &= g([0, 0]^T; f, \xi, \gamma, \mathbf{O}_i, \vec{n}_i)
 \end{aligned}$$

FIG. 9

Source image

Target image

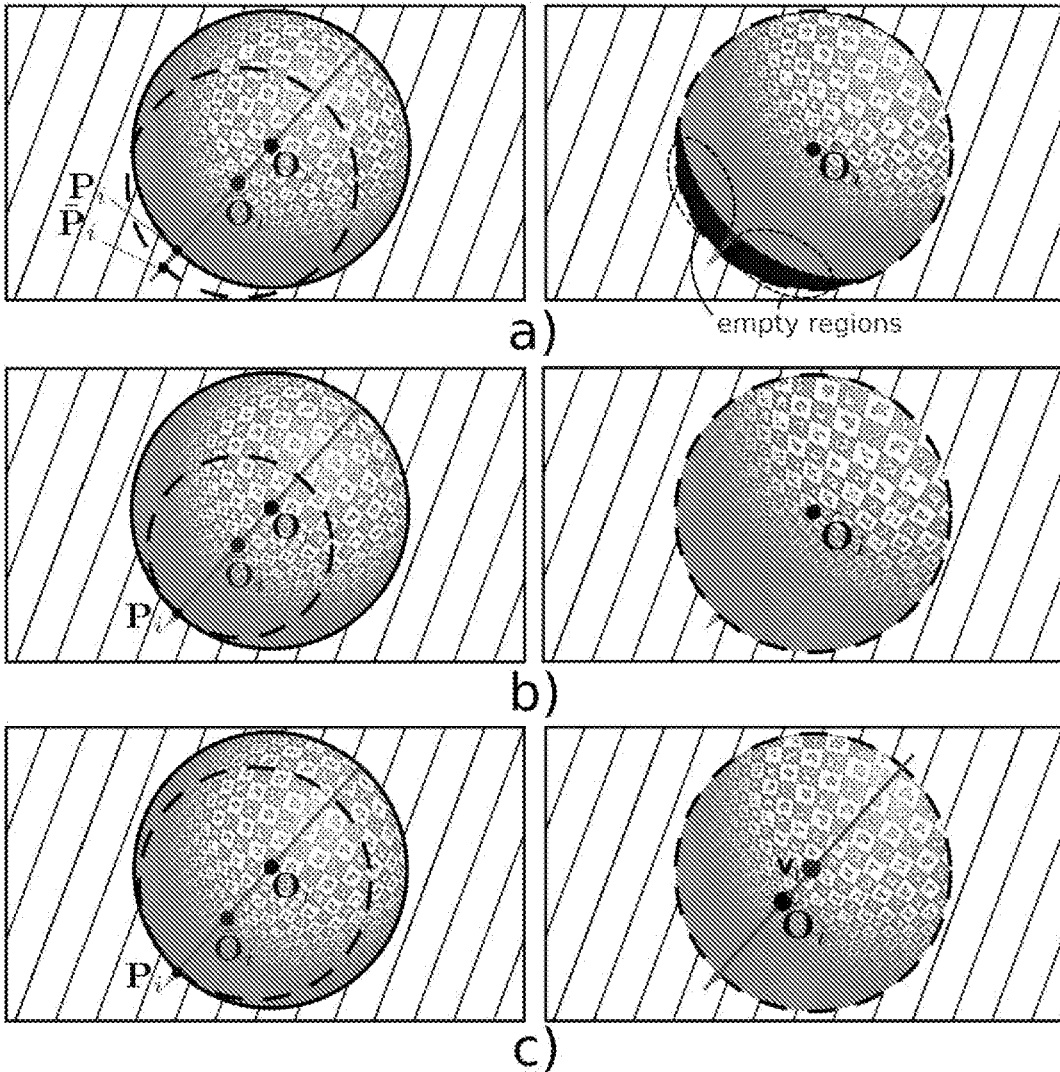


FIG. 10

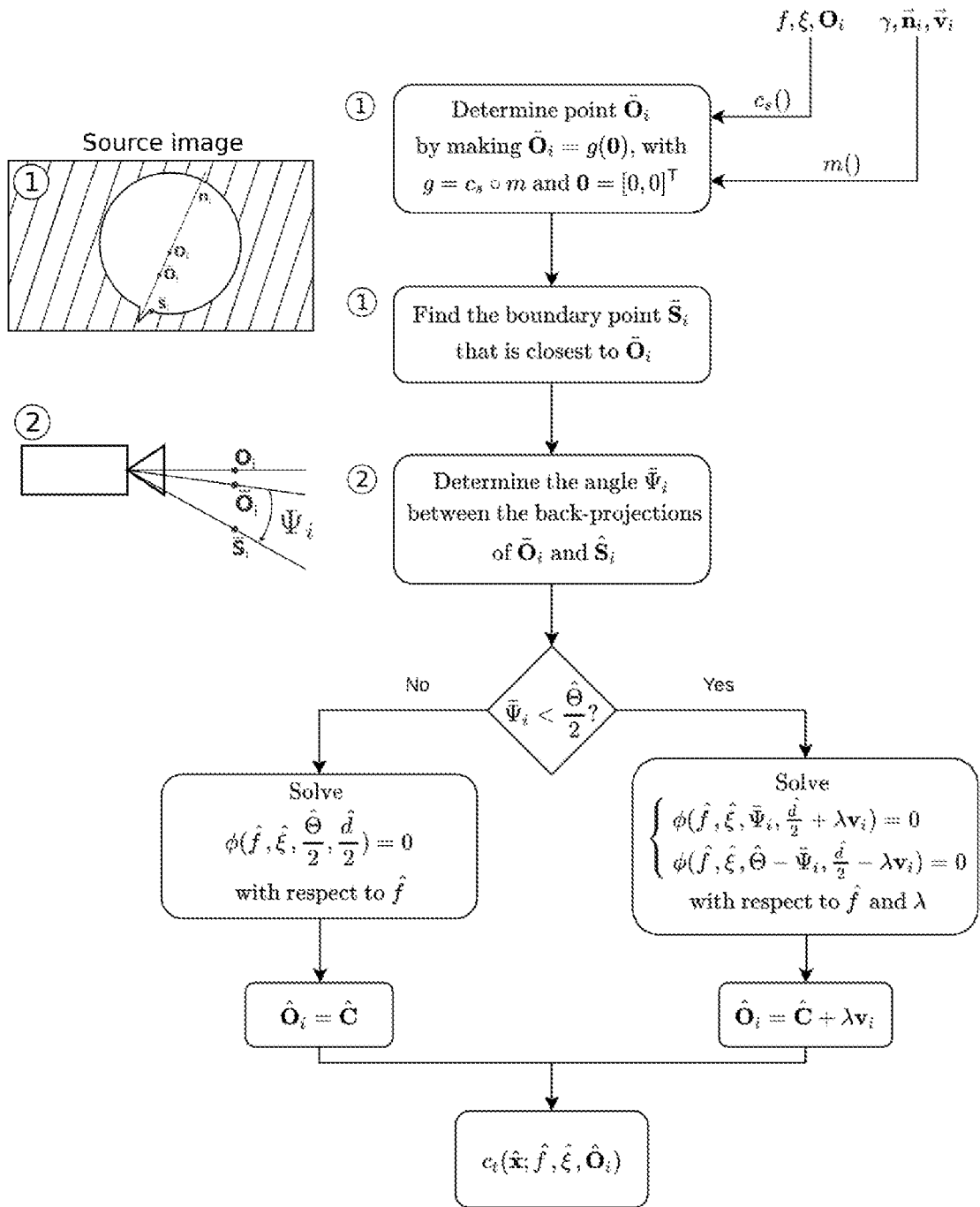


FIG. 11

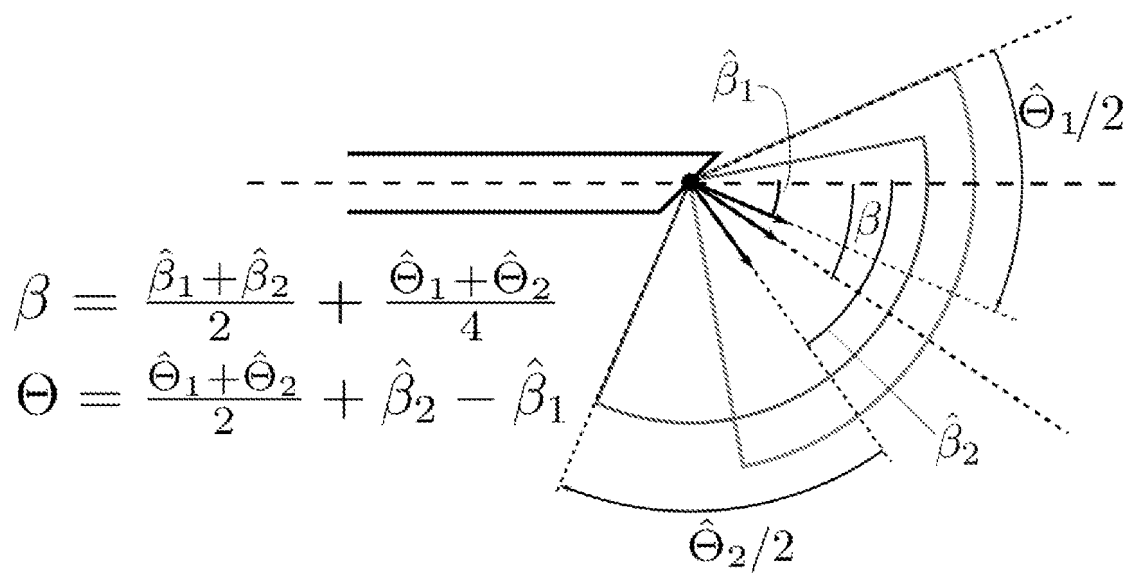
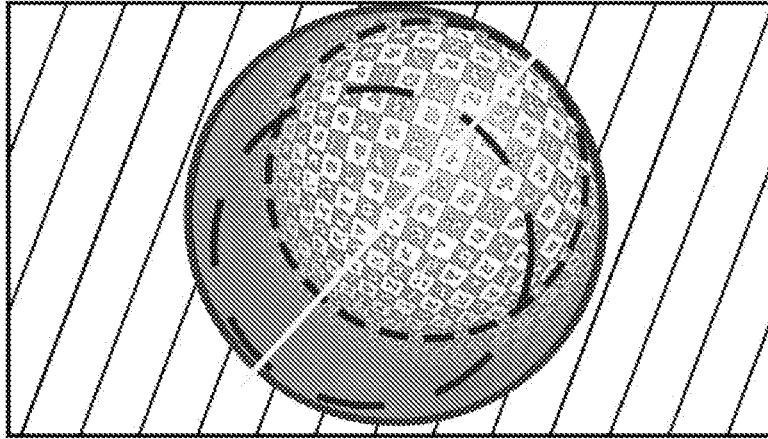
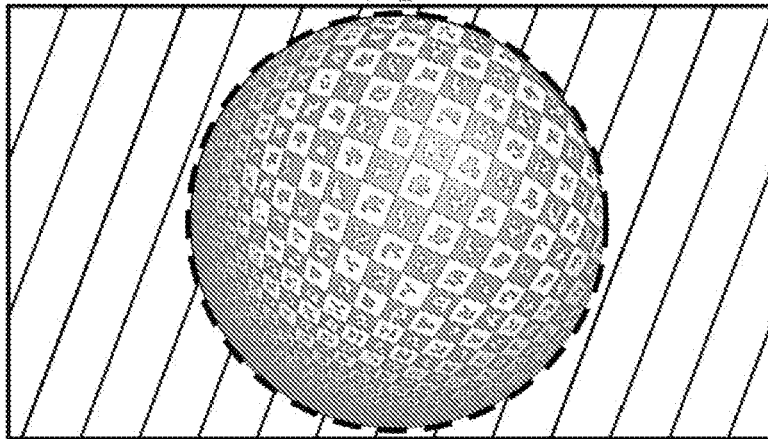


FIG. 12

Source image  
(lens cut  $\beta$ , FOV  $\Theta$ )



Target image  
(lens cut  $\hat{\beta}_1$ , FOV  $\hat{\Theta}_1$ )



Target image  
(lens cut  $\hat{\beta}_2$ , FOV  $\hat{\Theta}_2$ )

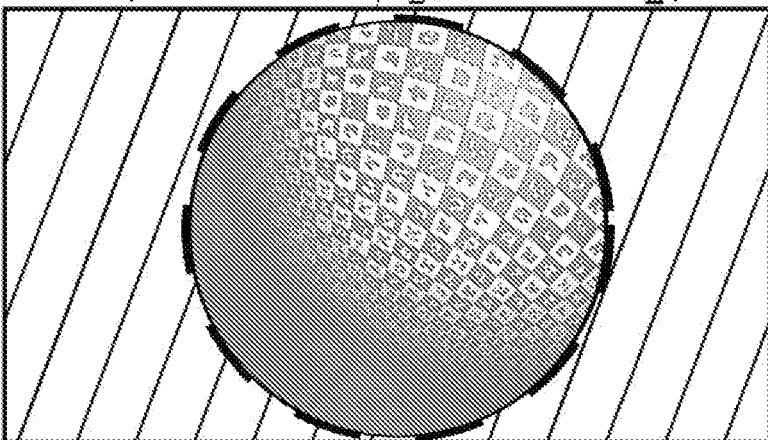


FIG. 13

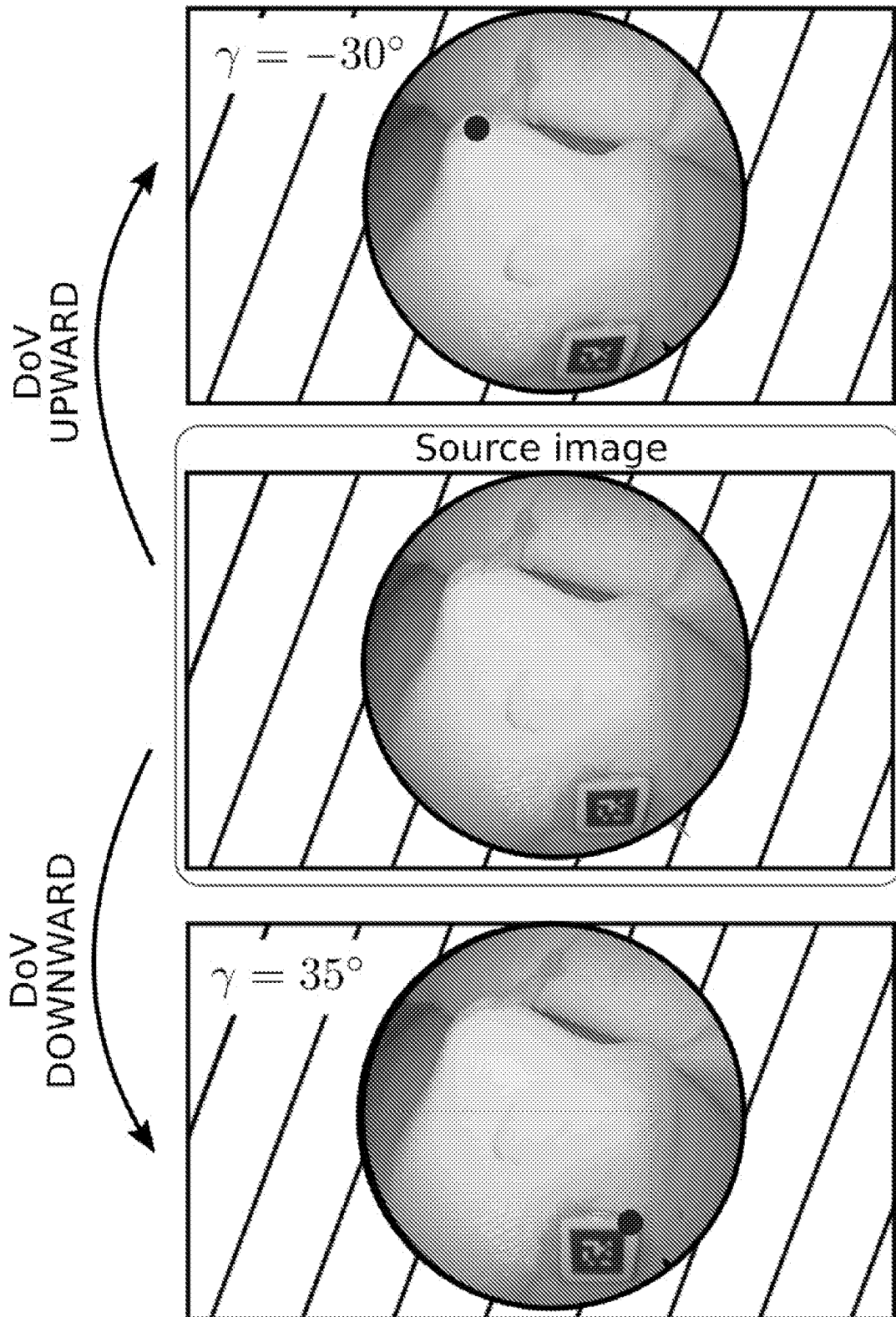


FIG. 14

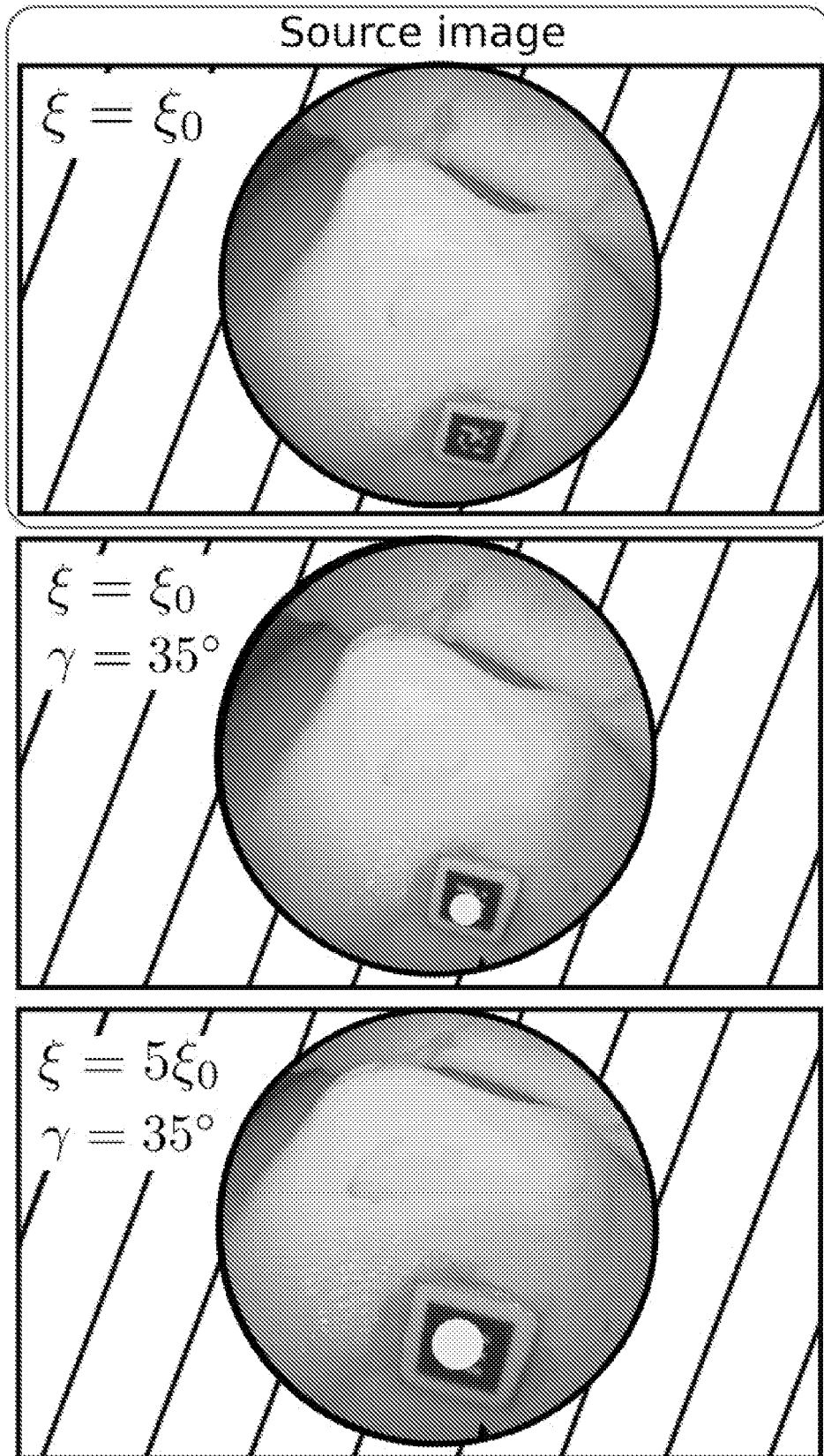


FIG. 15

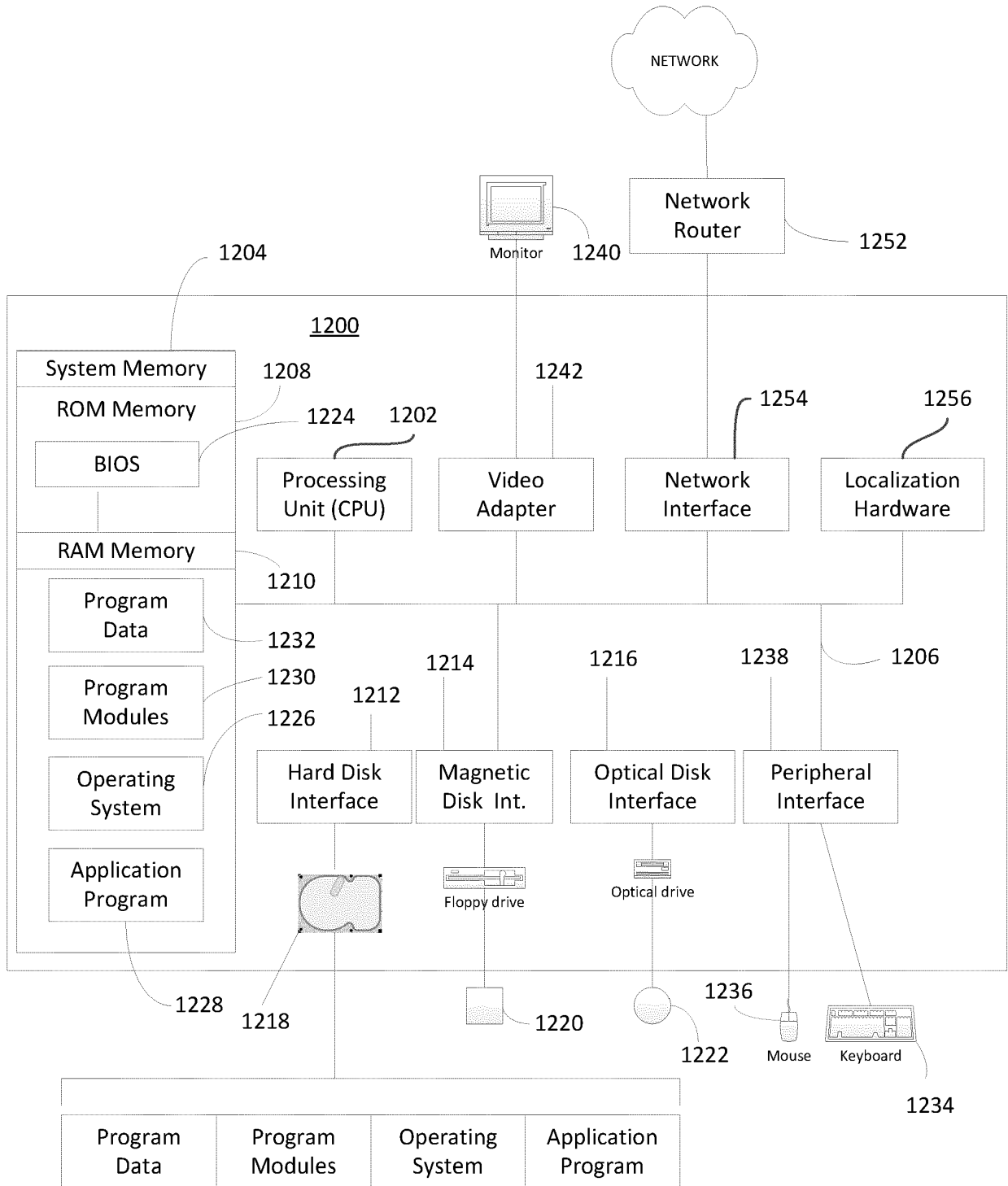


FIG. 16

**REFERENCES CITED IN THE DESCRIPTION**

*This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.*

**Patent documents cited in the description**

- US 62911986 [0001]
- US 62911950 [0001]
- US 20160209 A1 [0009]
- US 20050177026 A1 [0009]
- US 20120078049 A [0009]
- US 9307892 B [0009]
- US 005313306 A [0010]
- US 201540065799 A1 [0010]
- EP 3130276 A1 [0010]
- US 2014285676 A1, BARRETO [0011]
- US 62911950 B [0045] [0054]
- US 9438897 B [0053]
- US 14234907 B [0053]